University of California
Los Angeles

Part I: Steady States in Two-Species Particle Aggregation
Part II: Sparse Representations for Multiscale PDE

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mathematics

by

Alan Patrick Mackey

2015

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
MAR 2015		**00-00-2015 to 00-00-2015**

4. TITLE AND SUBTITLE
Part I: Steady States in Two-Species Particle Aggregation. Part II: Sparse Representations for Multiscale PDE

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California, Los Angeles,Department of Mathematics,Los Angeles,CA,90095

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)
11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The first part of this dissertation combines continuum limits of nonlocally interacting particles with stability analysis of nonlinear PDE to analyze the steady states of systems of pairwise-interacting particles. Models employing these assumptions cover a cornucopia of physical systems, from insect swarms and bacterial colonies to nanoparticle self-assembly. In this joint work with Theodore Kolokolnikov and Andrea Bertozzi [60], we study a continuum model with densities supported on co-dimension one curves for two-species particle interaction in R2, and apply linear stability analysis of concentric ring steady states to characterize the steady state patterns and instabilities which form. Conditions for linear well-posedness are determined and these results are compared to simulations of the discrete particle dynamics, showing predictive power of the linear theory. Part II continues the work started in [76], which proposes the sparse Fourier domain approximation of solutions to multiscale PDE problems by soft thresholding. In this joint work with Hayden Schaeffer and Stanley Osher [61], we show that the method enjoys a number of desirable numerical and analytic properties, including convergence for linear PDE and a modified equation resulting from the sparse approximation. We also extend the method to solve elliptic equations and introduce sparse approximation of differential operators in the Fourier domain. The effectiveness of the method is demonstrated on homogenization examples where its complexity is dependent only on the sparsity of the problem and constant in many cases.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**119**	

Abstract of the Dissertation

Part I: Steady States in Two-Species Particle Aggregation
Part II: Sparse Representations for Multiscale PDE

by

Alan Patrick Mackey

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2015

Professor Andrea Bertozzi, Co-chair

Professor Stanley J Osher, Co-chair

The first part of this dissertation combines continuum limits of nonlocally interacting particles with stability analysis of nonlinear PDE to analyze the steady states of systems of pairwise-interacting particles. Models employing these assumptions cover a cornucopia of physical systems, from insect swarms and bacterial colonies to nanoparticle self-assembly. In this joint work with Theodore Kolokolnikov and Andrea Bertozzi [60], we study a continuum model with densities supported on co-dimension one curves for two-species particle interaction in R^2, and apply linear stability analysis of concentric ring steady states to characterize the steady state patterns and instabilities which form. Conditions for linear well-posedness are determined and these results are compared to simulations of the discrete particle dynamics, showing predictive power of the linear theory.

Part II continues the work started in [76], which proposes the sparse Fourier domain approximation of solutions to multiscale PDE problems by soft thresholding. In this joint work with Hayden Schaeffer and Stanley Osher [61], we show that the method enjoys a number of desirable numerical and analytic properties,

including convergence for linear PDE and a modified equation resulting from the sparse approximation. We also extend the method to solve elliptic equations and introduce sparse approximation of differential operators in the Fourier domain. The effectiveness of the method is demonstrated on homogenization examples, where its complexity is dependent only on the sparsity of the problem and constant in many cases.

The dissertation of Alan Patrick Mackey is approved.

Alan Loddon Yuille

Luminita Aura Vese

Stanley J Osher, Committee Co-chair

Andrea Bertozzi, Committee Co-chair

University of California, Los Angeles

2015

To my family, friends, teachers, and collaborators.

The search for lost things is hindered by routine habits.

Table of Contents

List of Figures

Acknowledgments

I would like to thank my advisors Andrea Bertozzi and Stanley Osher for their guidance and support, which made possible all of the research presented in this dissertation. They have been an inspiration.

Andrea, thank you for your flexibility and encouragement that I pursue problems I loved, and for introducing me to Rick Chartrand, Brendt Wohlberg, Theodore Kolokolnikov, and so many others. Our first conversations about the aggregation equation are what convinced me to pursue applied math, and I am deeply grateful for that.

Stan, thank you for your complete tenacity in our sparse PDE work and for your patience with the many ideas I wanted to try. Our project introduced me to sparse representations, which became the largest piece of this dissertation and the greatest passion of my career.

I would also like to thank my other committee members, Luminita Vese and Alan Yuille, for their time, instruction, and helpful conversations. The extra time they spent with me and their discerning perspectives have been valuable gifts.

Other UCLA mathematicians have made these last five years a pleasure. Inwon Kim, Chris Anderson, Will Feldman, James von Brecht, Hui Sun, Marcus Roper, Joseph Teran, and Russell Caflisch have all supported me in their own ways and areas of expertise, for which I am very grateful.

I am indebted to my collaborators Theodore Kolokolnikov, Hayden Schaeffer, Rick Chartrand, Brendt Wohlberg, and Joseph Woodworth for the insights they've shared to make our work possible. I'm a better mathematician for working with each of them.

The research presented here was made possible in part by the National Science Foundation through grants DMS-0907931, EFRI-1024765, and CMMI-1435709. Additional funding was provided by UC Lab Fees Research Grant 12-LR-236660,

and ONR grants N00014-14-1-0444 and N00014-14-1-0683.

Vita

2006	Flinn Foundation Undergraduate Scholarship
2009	University of Arizona College of Science Galileo Circle Scholarship
2009	University of Arizona Honors College Undergraduate Research Grant
2010	University of Arizona Outstanding Senior in Mathematics
2010	B.A. Mathematics, University of Arizona, *Summa cum laude*
2011-2012	Graduate Teaching Assistant, UCLA
2012 Summer	Undergraduate Research Project Mentor, UCLA REU
2012-2015	Graduate Research Assistant, UCLA
2013 Summer	Research Assistant, Los Alamos National Laboratory
2014 Summer	Machine Learning Intern, Amazon.com, Inc.

Publications

Alan Mackey, Theodore Kolokolnikov, and Andrea Bertozzi. Two-species particle aggregation and stability of co-dimension one solutions. *Discrete and Continuous Dynamical Systems*, 34: 1411–1436, 2014.

Alan Mackey, Hayden Schaeffer, and Stanley Osher. On the Compressive Spectral Method. *Multiscale Modeling and Simlation*, 12(4): 1800–1827, 2014.

CHAPTER 1

Steady-States in Two-Species Particle Aggregation

1.1 Preface

This chapter is concerned with stability analysis of configurations formed by a large number of two species of pairwise-interacting particles in $\mathbb{R}^2$ [60]. The crux of our stability analysis comes down to the aggregation equation

$$u_t + \operatorname{div}(uv) = 0,$$
$$v = -\nabla K * u,$$

which is used to approximate the damped collective particle motion by the dynamics of a continuum mass $u(x)$ subject to the potential K.

To illustrate the mathematical tools used, we first introduce PDE limits for discrete systems, discuss the aggregation equation, give an example of the technique in action, and then proceed to the specifics of our project in section 1.2.

1.1.1 PDE Limits

A main step of this chapter is the use of a nonlocal PDE to describe aggregate particle interactions as the number of particles goes to infinity. Analysis of the system of particles is then shifted from high-dimensional coupled ODEs to a PDE. For our purposes, this continuum limit is partially an ansatz because we consider potentials for which the particles remain contained in a bounded set regardless of their number. Therefore, we assume from the start that the problem may

be formulated in terms of measures; for the case of finitely many particles, the corresponding measure is a sum of dirac masses. Once the correct PDE limit has been identified, it is often easiest to think of the case of finitely many particles as a special case of the equations for a general measure, i.e. the continuum limit. For more details, see sections 1.1.3 and 1.4.

Continuum and hydrodynamic limits appear in other contexts such as random interacting particle systems, statistical mechanics, and kinetic theory. In most of these areas, identification of the correct limit involves more complex scaling techniques and tools from probability theory. Discussion of these well-established topics is outside the scope of this dissertation, but interested readers may consult [13, 48, 74] and other references.

1.1.2 The Aggregation Equation

The continuum limit of the single-species version of the problem discussed in this chapter is given by the aggregation equation [53] in $\mathbb{R}^2$:

$$u_t + \operatorname{div}(uv) = 0,$$
$$v = -\nabla K * u,$$

which models a scalar quantity $u(x,t)$, typically referred to as "mass", advected by a velocity field $v(x,t)$. The velocity field arises as the gradient of a potential K, typically playing the role of some kind of gravitation, convolved with u. Thus, the aggregation equation is commonly used to model the overdamped dynamics of a quantity with nonlocal attraction/repulsion specified by K. It appears in models of bacterial colonies, swarms, robotic control, and physical chemistry. For more applications and models, see section 1.2.

Due to the advective nature of the aggregation equation, its analysis (see [7, 53, 5] and many others) has much in common with that of the fluid Euler equation [62]. In the sense of the Helmholtz Decomposition for vector fields, the

Euler and aggregation equations are orthogonal: in the Euler equation the velocity field is divergence-free, while in the aggregation equation it is the gradient of a potential.

To illustrate this point, consider the aggregation equation with the Newtonian potential:

$$u_t + \operatorname{div}(uv) = 0,$$
$$v = \nabla \Delta^{-1} u = \nabla N * u,$$

i.e. $v = \nabla \phi$ where $\Delta \phi = \omega$. The Poisson equation for ϕ is solved by convolution with the Newtonian potential $N(x) = \frac{1}{2\pi} \log |x|$. For more about aggregation in this context, see [6].

Now recall that the vorticity form of the two-dimensional Euler equation is [62]

$$\omega_t + \operatorname{div}(\omega v) = 0$$
$$v = \nabla^\perp \Delta^{-1} \omega = \nabla^\perp N * \omega,$$

i.e. $v = \nabla^\perp \psi$ where $\Delta \psi = \omega$. Up to the perpendicular gradient, the two equations are identical. For the vorticity equation, the convolution defining the velocity is known as the Biot-Savart law [62]:

$$v(x,t) = \int_{\mathbb{R}^d} K_2(x-y)\omega(y,t)\, dy$$

where

$$K_2 = \nabla^\perp N = \frac{1}{2\pi}\left(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2}\right)^T.$$

For the two-species problem discussed in this chapter, the continuum limit is a generalization of the aggregation equation to two quantities with different interactions. Detailed knowledge of the usual, single-quantity aggregation equation is not a prerequisite for our analysis here, but many of the steps in section 1.4 will be familiar to readers with a background in aggregation or other advective PDE. Sections 1.4-1.6 closely follow [87].

1.1.3 An Example

To introduce the continuum limit technique in our context, we consider the limit of a problem with a single type of particle. Assume the particles occupy positions $\mathbf{x}_1 \dots \mathbf{x}_N \in \mathbb{R}^d$, and that every pair of particles (i, j) in the absence of others will move to minimize the potential energy

$$P(|\mathbf{x}_i - \mathbf{x}_j|)$$

where $P(x)$ is a function with a unique minimum for $x > 0$. With all particles present, they jointly move to minimize the total potential energy

$$E(\mathbf{x}_1, \dots, \mathbf{x}_N) \sum_{1 \le i < j \le N} P(|\mathbf{x}_i - \mathbf{x}_j|). \tag{1.1}$$

To approach the problem, we consider the continuum limit of (1.1):

$$E_c(u) = \frac{1}{2} \int_x \int_y P(|x - y|) u(x) u(y) \, dxdy. \tag{1.2}$$

In the above energy, $u(x)$ represents a continuum mass in space. If $u(x) = \sum_{i=1}^{N} \delta_{\mathbf{x}_i}(x)$, it is easily verified that (1.2) reduces to (1.1) with the assumption that particles do not interact with themselves.

If we assume further that $u(x, t)$ evolves the a gradient flow of the energy (1.2), we arrive at the aggregation equation

$$u_t + \operatorname{div}(uv) = 0,$$
$$v = -\nabla P * u,$$

introduced above. See [4] for details.

Alternatively, we can reach the continuum limit from the discrete energy E directly. This is the approach taken in sections 1.3 and 1.4 below.

1.2 Introduction

The collective behavior of systems of interacting particles gives rise to emergent phenomena in physics, biology, chemistry, and other disciplines. Models of pairwise-interacting agents find applications in the biological contexts of locust swarms [3, 83, 82], animal flocks [25, 55, 63], and bacterial colonies [85]. These mathematical approaches to swarming have also inspired algorithms for cooperative control of robotic vehicles [57]. More questions for nonlocal particle systems arise in physical chemistry: the self-assembly of nanoparticles [23, 44] and arrangement of ions into spheres [58, 59] are just two examples. In the physical contexts of granular gasses [71] and molecular dynamics simulations of matter [42], particle systems also have a central role.

All of the above models, however multifaceted, share the same footing. Some number of particles interact with each other pairwise such that any two particles will repel each other when they are close and attract when they are far; typically, this attractive force disappears at very long distances. These interactions can generate rich steady states relevant to the models in which they arise.

Consider the case when the forces arise due to a pairwise interaction energy

$$E(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_{i \neq j} P(|\mathbf{x}_i - \mathbf{x}_j|)$$

where $\mathbf{x}_i$ denotes the position in $\mathbb{R}^d$ of the i^{th} particle and $P(r)$ is the potential energy between two particles. $P(r)$ is usually a function with a unique minimum such that the force on one particle due to another, $F(r) := -P'(r)$, enjoys the repulsive-attractive properties mentioned above. In this framework, a steady state pattern can be understood as a minimizer of E.

We call the potential E *confining* if its minimizing configurations $\mathbf{x}_1, \ldots, \mathbf{x}_N$ stay contained inside a compact set as $N \to \infty$. The question of whether or not a given function P will result in a confining potential has been addressed in

terms of the notion of H-stability in statistical mechanics; see [32]. For confining potentials, particles in ground states may reside in space-filling, co-dimension zero configurations or concentrate on co-dimension one manifolds. The question of which occurs is answered in [2] and the problem of characterizing ground states (or steady states) has been discussed in [50, 51, 79, 87], and elsewhere. Applications where both co-dimension zero and one solutions are of importance include bacterial colony growth under stress, point vortex theory, and the Thomson problem [1, 12, 22, 85].

It is a natural extension of the above work to consider the analogous problems for two particle species; i.e., when more than one type of particle is present in the interactions. Two-species models are relevant for the phenomena observed in [59], where two types of macroions in solution will self-recognize and assemble into hollow spherical structures. This self-recognition of particle species is a robust phenomenon observed in many of the numerical experiments considered in this chapter. Two-species models also find application in large scale pedestrian movement [73], and the well-posedness of said models has been considered in [27]; a general treatment of well-posedness for the two-species problem is given in [29]. Other applications include opinion formation in groups consisting of ordinary individuals and strong leaders [33] and two-species group consensus [36]. Two-species bacterial aggregation driven by chemotaxis and diffusion is another area of active research, where [54] employs a two-species model for localized vortex formation in bacterial colonies. Global existence and finite time blowup are considered in [24] and [37], [90] treats the n-species problem, and [45] and [80] discuss the stability of uniform density and homogenous steady states.

Our numerical experiments have revealed phenomena which did not appear in the single species problem. Particle species either mix or segregate based on the relative strengths of the inter-species and intra-species forces, and occasionally settle in domains with irregular boundaries including cusps. Asymmetric steady

states (which represent local minimizers of the potential) can be observed, and nontrivial structures form when particles are H-stable. These and other features indicate a substantial increase in complexity of the two-species problem over the single species problem.

In this chapter, our objective is to characterize steady states formed in the two-species aggregation problem in the absence of diffusion. Inspired by physical [58] and numerical experiments exhibiting steady states supported on or near co-dimension one manifolds, we wish to determine the circumstances under which these steady states form and characterize their properties when possible. For example, the authors of [87] were able to determine when the steady states exhibited three-fold or five-fold symmetry, say.

One approach to this problem would be to work directly with the system of $N_1 + N_2$ coupled ODEs, where N_1 is the number of particles of species I and N_2 is analogous. The primary difficulty is that the number of unknowns increases as the number of particles increases. It is still possible to pursue linear stability analysis, but the number of linearly independent perturbations to consider grows with the number of particles. Additionally, linear stability analysis hinges upon finding a meaningful basis of eigen-perturbations (or modes) of the system. This reduces to an eigenvector problem, but meaningful interpretation of the instabilities becomes difficult. Moreover, the results will apply only for a particular choice of N_1 and N_2, even though from physical and numerical experiemnts we expect that in many cases the nature of the instabilities will not change after a certain number of particles has been reached. For example, a mode-three instability manifesting as a triangular arrangement of particles in the steady state (as in [51]) persists even as more particles are considered.

An alternative method is to consider the limit as $N_1, N_2 \to \infty$. For the single-species problem, this was the approach used successfully in [51, 87]. The method of considering a *continuum limit* or *hydrodynamic limit* such as this has roots in

statistical mechanics, kinetic theory, and fluids [13, 48, 74]. When the potential governing the particle interactions is confining, the continuum limit is meaningful. When the potential is not confining for the single-species case, the particles arrange into a regular lattice. In the two-species case, however, the lattice structure formed may have nontrivial structure depending on the relative interaction strengths between the two types of particles. This phenomenon is explored numerically in section 1.8.

In this part, we study the continuum limit of the problem and the stability of a steady state consisting of two concentric rings of constant density. The theory put forth accurately predicts instabilities observed in numerical experiments and the breakup of ring solutions into fully two-dimensional patterns. The arguments presented here are restricted to the two-dimensional problem, but adapting the theory of [87] could generalize the results to higher dimensions.

1.3 Problem Description

Consider two species of particles, type I and type II, which occupy positions $\mathbf{x}_1(t), \ldots, \mathbf{x}_{N_1}(t), \mathbf{y}_1(t), \ldots, \mathbf{y}_{N_2}(t)$ in $\mathbb{R}^2$. Steady state patterns are minimizers of the pairwise interaction energy

$$\begin{aligned} &E(\mathbf{x}_1, \ldots, \mathbf{x}_{N_1}, \mathbf{y}_1, \ldots, \mathbf{y}_{N_2}) \qquad (1.3)\\ &= \sum_{\substack{i,j=1\\ i\neq j}}^{N_1} P_1(|\mathbf{x}_i - \mathbf{x}_j|) + \sum_{\substack{i,j=1\\ i\neq j}}^{N_2} P_2(|\mathbf{y}_i - \mathbf{y}_j|) + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} P_3(|\mathbf{x}_i - \mathbf{y}_j|)\\ &=: \sum_{\substack{i,j=1\\ i\neq j}}^{N_1} V_1\left(\frac{1}{2}|\mathbf{x}_i - \mathbf{x}_j|^2\right) + \sum_{\substack{i,j=1\\ i\neq j}}^{N_2} V_2\left(\frac{1}{2}|\mathbf{y}_i - \mathbf{y}_j|^2\right) + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} V_3\left(\frac{1}{2}|\mathbf{x}_i - \mathbf{y}_j|^2\right). \end{aligned}$$

Above, $V_i(s) := P_i(\sqrt{2s})$, $i = 1, 2, 3$, is simply a change of variables to simplify the calculations.

We are interested in gradient flow equations associated with (1.3):

$$\begin{aligned}\frac{d\mathbf{x}_i}{dt} &= -\frac{\partial E}{\partial \mathbf{x}_i} = -\nabla_{\mathbf{x}_i} E \\ &= \frac{1}{N_1}\sum_{\substack{j=1\\ j\neq i}}^{N_1} g_1\left(\frac{1}{2}|\mathbf{x}_i-\mathbf{x}_j|^2\right)(\mathbf{x}_i-\mathbf{x}_j) + \frac{1}{N_1}\sum_{j=1}^{N_2} g_3\left(\frac{1}{2}|\mathbf{x}_i-\mathbf{y}_j|^2\right)(\mathbf{x}_i-\mathbf{y}_j) \\ &= \frac{1}{N_1}\sum_{\substack{j=1\\ j\neq i}}^{N_1} g_1\left(\frac{1}{2}|\mathbf{x}_i-\mathbf{x}_j|^2\right)(\mathbf{x}_i-\mathbf{x}_j) + \mu\frac{1}{N_2}\sum_{j=1}^{N_2} g_3\left(\frac{1}{2}|\mathbf{x}_i-\mathbf{y}_j|^2\right)(\mathbf{x}_i-\mathbf{y}_j)\end{aligned} \tag{1.4}$$

for $i = 1, \ldots, N_1$, and

$$\begin{aligned}\frac{d\mathbf{y}_i}{dt} &= -\frac{\partial E}{\partial \mathbf{y}_i} = -\nabla_{\mathbf{y}_i} E \\ &= \frac{1}{N_1}\sum_{\substack{j=1\\ j\neq i}}^{N_2} g_2\left(\frac{1}{2}|\mathbf{y}_i-\mathbf{y}_j|^2\right)(\mathbf{y}_i-\mathbf{y}_j) + \frac{1}{N_1}\sum_{j=1}^{N_1} g_3\left(\frac{1}{2}|\mathbf{y}_i-\mathbf{x}_j|^2\right)(\mathbf{y}_i-\mathbf{x}_j) \\ &= \mu\frac{1}{N_2}\sum_{\substack{j=1\\ j\neq i}}^{N_2} g_2\left(\frac{1}{2}|\mathbf{y}_i-\mathbf{y}_j|^2\right)(\mathbf{y}_i-\mathbf{y}_j) + \frac{1}{N_1}\sum_{j=1}^{N_1} g_3\left(\frac{1}{2}|\mathbf{y}_i-\mathbf{x}_j|^2\right)(\mathbf{y}_i-\mathbf{x}_j)\end{aligned} \tag{1.5}$$

for $j = 1, \ldots, N_2$. The right-hand sides of (1.4) and (1.5) have been divided by N_1 as a simple rescaling of time, and in the second line the parameter $\mu := N_2/N_1$ has been introduced. The factors $1/N_1$ and $1/N_2$ may also be seen as normalizing each species by its total particle number or 'mass', in which case μ represents the relative mass of species II to species I.

In the above, $g_i(s) := -\frac{dV_i}{ds}(s)$ for $i = 1, 2, 3$ give the 'forces' due to the potentials. Note that $g_i(s)$ is the derivative of the rescaled potential V with respect to its argument $s = \frac{1}{2}r^2$, where r represents true particle distance. As such, $g_i(s)$ represents the force only with respect to the rescaled space variable $\frac{1}{2}r^2$. The true physical force— the derivative of the potential with respect to true particle distances r and not just with respect to its argument $\frac{1}{2}r^2$— has magnitude $rg_i(\frac{1}{2}r^2)$. The difference is illustrated in figure 1.3.

One can think of the gradient flow either as an approximation to overdamped second order physical dynamics or simply as a means to identify minimizers of the energy. In the next section, we show that for large numbers of particles the gradient flow system (1.4), (1.5) may be approximated by a nonlocal PDE system of advection equations similar to the Birkhoff-Rott equation for vortex sheets (c.f. [62, 78]), for which linear stability is reduced to a sequence of eigenvalue problems. Criteria for the stability of each element in a basis of perturbations, and for linear well-posedness of the concentric ring solution, are derived. Numerical examples are presented, which demonstrate strong agreement with the theory put forth.

In this work we consider the following potentials, which have all been considered in the literature for the single species problem [32, 51, 56, 87]: the Morse potential

$$V_i(s) = C_{r_i} e^{-\sqrt{2s}/l_{r_i}} - C_{a_i} e^{-\sqrt{2s}/l_{a_i}},$$

power law forces

$$g_i(s) = s^{p_i} - s^{q_i},$$

and smoothed step discontinuity forces

$$g_i(s) = \frac{\tanh\big[a_i(1-\sqrt{2s})\big] + b_i}{\sqrt{2s}}$$

with steady states pictured in figure 1.1. Combinations of all three of the above types (and others) are also plausible; for example, g_1 could arise from a power law, g_2 from the Morse potential, and g_3 from the tanh force. See figure 1.2.

1.4 The Continuum Limit

For the two-species case, we will say that an energy E such as (1.3) is *confining* if its minimizing configurations $\mathbf{x}_1, \ldots, \mathbf{x}_{N_1}$, $\mathbf{y_1}, \ldots, \mathbf{y}_{N_2}$ stay contained inside a compact set as N_1 and $N_2 \to \infty$. Under the assumption that the energy E

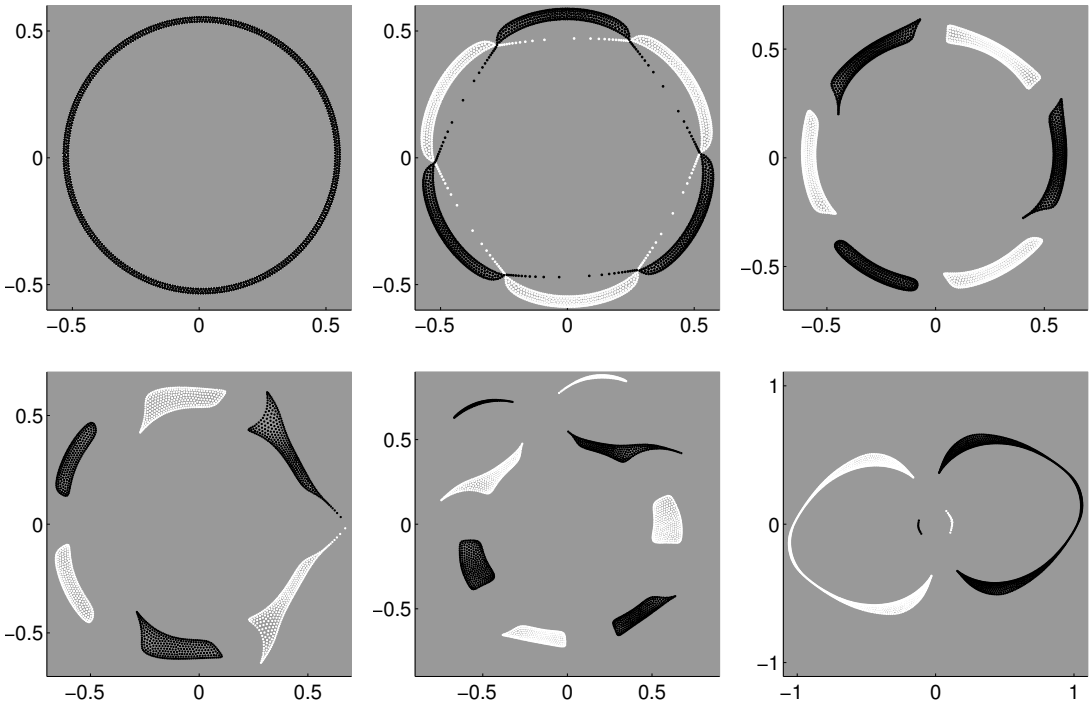


Figure 1.1: Steady states of the tanh potential with $\mu = 1$, $a_1 = a_2 = 10$, $b_1 = b_2 = 0.1$, $a_3 = 2$ and $b_3 = -0.7, -0.5, -0.3, -0.1$, 0.1, and 0.3 (from top left to bottom right). Each steady state consists of 1000 white particles and 1000 black, with (1.4), (1.5) evolved to a final time $t = 2000$.

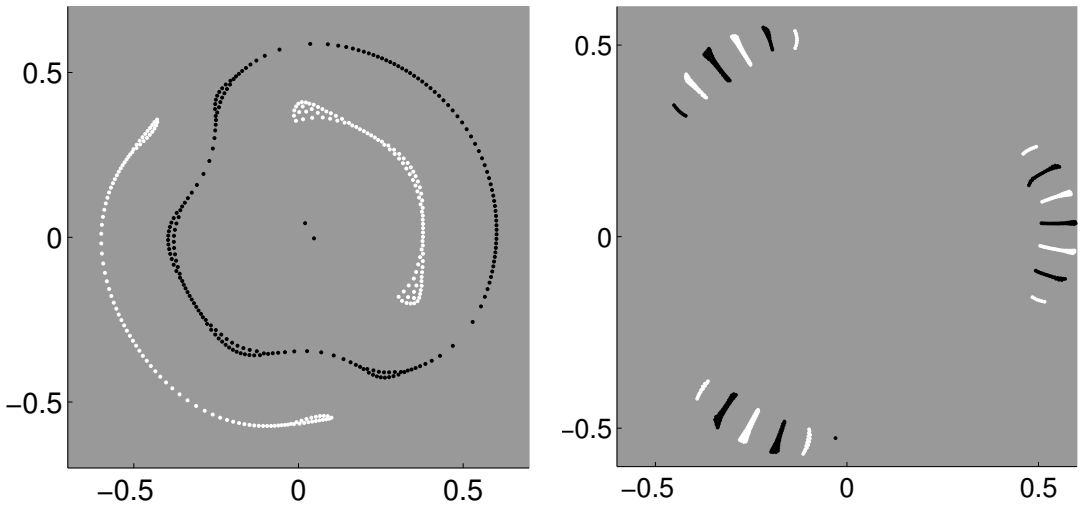


Figure 1.2: Steady states with g_1 and g_2 tanh forces and g_3 a Morse force. On the left, $a_1 = a_2 = 10$, $b_1 = b_2 = -0.3$, $C_a = l_a = C_r = 1$ and $l_r = 0.1$. On the right, $a_1 = a_2 = 10$, $b_1 = b_2 = 0.1$, $C_a = l_a = C_r = 1$ and $l_r = 0.02$.

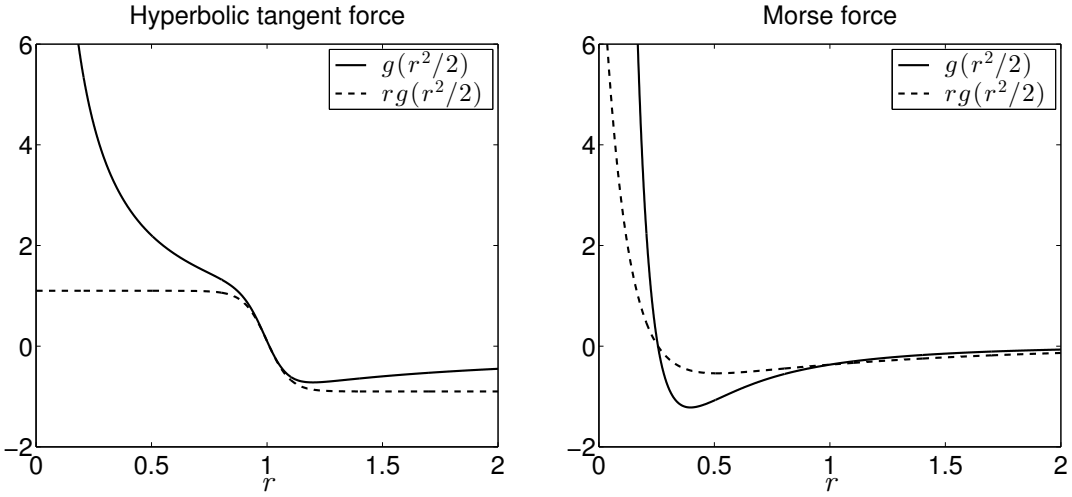


Figure 1.3: Hyperbolic tangent (tanh) and Morse forces $g\left(\frac{1}{2}r^2\right)$ with respect to rescaled space, and $rg\left(\frac{1}{2}r^2\right)$ with respect to physical space. Tanh parameters $a = 10, b = 0.1$; Morse parameters $C_a = l_a = C_r = 1$, $l_r = 0.1$.

is confining, the configurations of discrete particles approach continuum spatial densities ρ_1 and ρ_2.

As we are interested in the stability of concentric ring solutions, we seek densities which are supported along one-dimensional curves $\Gamma_1(t) = \mathbf{\Phi}_1(\alpha, t)$ and $\Gamma_2(t) = \mathbf{\Phi}_2(\alpha, t)$ (parameterized by $\alpha \in D \subset \mathbb{R}$) which evolve with velocity fields $\mathbf{v}_1$ and $\mathbf{v}_2$; that is,

$$\begin{aligned} \frac{\partial \mathbf{\Phi}_1}{\partial t}(\alpha, t) &= \mathbf{v}_1(\mathbf{\Phi}_1(\alpha, t), t) \\ \frac{\partial \mathbf{\Phi}_2}{\partial t}(\alpha, t) &= \mathbf{v}_2(\mathbf{\Phi}_2(\alpha, t), t). \end{aligned} \tag{1.6}$$

The velocity fields $\mathbf{v}_1$ and $\mathbf{v}_2$ are determined from the respective densities by the continuum limits of equations (1.4) and (1.5): for $\mathbf{x} \in \mathbb{R}^2$,

$$\begin{aligned} \mathbf{v}_1(\mathbf{x}, t) &= \int_{\mathbb{R}^2} g_1\left(\frac{1}{2}|\mathbf{x}-\mathbf{y}|^2\right)(\mathbf{x}-\mathbf{y})\rho_1(\mathbf{y}, t) + g_3\left(\frac{1}{2}|\mathbf{x}-\mathbf{y}|^2\right)(\mathbf{x}-\mathbf{y})\rho_2(\mathbf{y}, t)\, d\mathbf{y} \\ \mathbf{v}_2(\mathbf{x}, t) &= \int_{\mathbb{R}^2} g_2\left(\frac{1}{2}|\mathbf{x}-\mathbf{y}|^2\right)(\mathbf{x}-\mathbf{y})\rho_2(\mathbf{y}, t) + g_3\left(\frac{1}{2}|\mathbf{x}-\mathbf{y}|^2\right)(\mathbf{x}-\mathbf{y})\rho_1(\mathbf{y}, t)\, d\mathbf{y}, \end{aligned} \tag{1.7}$$

where we must assume that $N_2/N_1 \to \mu$ as $N_1, N_2 \to \infty$. The parameter μ is absorbed into ρ_2 and still represents the relative mass

$$\mu = \frac{\int_{\mathbb{R}^2} \rho_2}{\int_{\mathbb{R}^2} \rho_1}.$$

To determine the dynamics of ρ_1, ρ_2, $\mathbf{\Phi}_1$, and $\mathbf{\Phi}_2$ completely, we impose conservation of mass:

$$\begin{aligned} \frac{\partial \rho_1}{\partial t} + \nabla\cdot(\rho_1 \mathbf{v}_1) &= 0 \\ \frac{\partial \rho_2}{\partial t} + \nabla\cdot(\rho_2 \mathbf{v}_2) &= 0, \end{aligned} \tag{1.8}$$

which is implicit in the particle formulation of the problem. Equations (1.6), (1.7), and (1.8) specify a nonlocal coupled advection system determining $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$, from which ρ_1 and ρ_2 will be recovered later; see, for example, [78] and [87].

It is worth pointing out here that the ρ_i are densities of measures singular with respect to Lebesgue measure on $\mathbb{R}^2$. Therefore, they should solve (1.8) weakly, or in the sense of distributions. We will assume there exist f_i locally integrable on $\mathbb{R}$ such that for Borel sets $E \subseteq \mathbb{R}^2$,

$$\begin{aligned}\int_E \rho_i(\mathbf{x},t)\,d\mathbf{x} &= \int\limits_{\{\alpha:\,\mathbf{\Phi}_i(\alpha,t)\in E\}} f_i(\alpha,t)\,d\alpha \\ &=: \int\limits_{\{\alpha:\,\mathbf{\Phi}_i(\alpha,t)\in E\}} \rho_i^s(\alpha,t)\left|\frac{\partial \mathbf{\Phi}_i}{\partial \alpha}\right| d\alpha\end{aligned}$$

where ρ_i^s admits the natural interpretation of the density along the surface Γ_i. It then follows that for $\psi \in C_c^\infty(\mathbb{R}^2 \times [0,\infty])$,

$$\begin{aligned}\int_0^\infty \int_{\mathbb{R}^2} \psi(\mathbf{x},t)\rho_i(\mathbf{x},t)\,d\mathbf{x}dt &= \int_0^\infty \int_D \psi(\mathbf{\Phi}_i(\alpha,t),t) f_i(\alpha,t)\,d\alpha dt \\ &= \int_0^\infty \int_D \psi(\mathbf{\Phi}_i(\alpha,t),t)\rho_i^s(\alpha,t)\left|\frac{\partial \mathbf{\Phi}_i}{\partial \alpha}\right| d\alpha dt.\end{aligned}$$

One can now integrate by parts from (1.8) to define what it means for ρ_i to be a solution: for all $\psi \in C_c^\infty(\mathbb{R}^2 \times [0,\infty])$,

$$\int_0^\infty \int_D \left(\frac{\partial \psi}{\partial t} + \mathbf{v}_i \cdot \nabla \psi\right)(\mathbf{\Phi}_i(\alpha,t),t) f_i(\alpha,t)\,d\alpha dt = 0.$$

Noting that $\left(\frac{\partial \psi}{\partial t} + \mathbf{v}_i \cdot \nabla \psi\right)(\mathbf{\Phi}_i(\alpha,t),t) = \frac{d}{dt}\psi(\mathbf{\Phi}_i(\alpha,t),t))$, one can integrate by parts to get

$$\begin{aligned}0 &= \int_0^\infty \int_D \frac{d}{dt}\psi(\mathbf{\Phi}_i(\alpha,t),t)) f_i(\alpha,t)\,d\alpha dt \\ &= 0 - \int_0^\infty \int_D \psi(\mathbf{\Phi}_i(\alpha,t),t))\frac{\partial}{\partial t} f_i(\alpha,t)\,d\alpha dt,\end{aligned}$$

where the boundary term drops out because ψ is compactly supported. It follows that $f(\alpha,t) \equiv f(\alpha,0) =: f^0(\alpha)$.

We also rewrite (1.7) in terms of integrals along Γ_1 and Γ_2:

$$\begin{aligned}\mathbf{v}_1(\mathbf{x},t) = \int_D &g_1\left(\frac{1}{2}|\mathbf{x}-\mathbf{\Phi}_1(\alpha,t)|^2\right)(\mathbf{x}-\mathbf{\Phi}_1(\alpha,t)) f_1^0(\alpha) + \\ &g_3\left(\frac{1}{2}|\mathbf{x}-\mathbf{\Phi}_2(\alpha,t)|^2\right)(\mathbf{x}-\mathbf{\Phi}_2(\alpha,t)) f_2^0(\alpha)\,d\alpha,\end{aligned}$$

$$\mathbf{v}_2(\mathbf{x},t) = \int_D g_2\left(\frac{1}{2}|\mathbf{x}-\mathbf{\Phi}_2(\alpha,t)|^2\right)(\mathbf{x}-\mathbf{\Phi}_2(\alpha,t))f_2^0(\alpha) + g_3\left(\frac{1}{2}|\mathbf{x}-\mathbf{\Phi}_1(\alpha,t)|^2\right)(\mathbf{x}-\mathbf{\Phi}_1(\alpha,t))f_1^0(\alpha)\,d\alpha.$$

Appealing to (1.6) then yields

$$\begin{aligned}\frac{\partial \mathbf{\Phi}_1}{\partial t}(\alpha,t) &= \mathbf{v}_1(\mathbf{\Phi}_1(\alpha,t),t)\\ &= \int_D g_1\left(\frac{1}{2}|\mathbf{\Phi}_1(\alpha,t)-\mathbf{\Phi}_1(\alpha',t)|^2\right)(\mathbf{\Phi}_1(\alpha,t)-\mathbf{\Phi}_1(\alpha',t))f_1^0(\alpha')\\ &\quad + g_3\left(\frac{1}{2}|\mathbf{\Phi}_1(\alpha,t)-\mathbf{\Phi}_2(\alpha',t)|^2\right)(\mathbf{\Phi}_1(\alpha,t)-\mathbf{\Phi}_2(\alpha',t))f_2^0(\alpha')\,d\alpha'\end{aligned} \tag{1.9}$$

and

$$\begin{aligned}\frac{\partial \mathbf{\Phi}_2}{\partial t}(\alpha,t) &= \mathbf{v}_2(\mathbf{\Phi}_2(\alpha,t),t)\\ &= \int_D g_2\left(\frac{1}{2}|\mathbf{\Phi}_2(\alpha,t)-\mathbf{\Phi}_2(\alpha',t)|^2\right)(\mathbf{\Phi}_2(\alpha,t)-\mathbf{\Phi}_2(\alpha',t))f_2^0(\alpha')\\ &\quad + g_3\left(\frac{1}{2}|\mathbf{\Phi}_2(\alpha,t)-\mathbf{\Phi}_1(\alpha',t)|^2\right)(\mathbf{\Phi}_2(\alpha,t)-\mathbf{\Phi}_1(\alpha',t))f_1^0(\alpha')\,d\alpha'.\end{aligned} \tag{1.10}$$

The two equations above determine $\mathbf{\Phi}_i$. With these in hand, all that is left is to determine ρ_i; for this,

$$0 = \frac{\partial}{\partial t}f_i(\alpha,t) = \frac{\partial}{\partial t}\left(\rho_i^s(\alpha,t)\left|\frac{\partial \mathbf{\Phi}_i}{\partial \alpha}\right|\right)$$

implies that

$$\frac{\partial \rho_i^s}{\partial t} = -\rho_i^s \frac{\frac{\partial}{\partial t}\left|\frac{\partial \mathbf{\Phi}_i}{\partial \alpha}\right|}{\left|\frac{\partial \mathbf{\Phi}_i}{\partial \alpha}\right|},$$

which is enough.

Note that when $\rho_1(\mathbf{x},t) = \sum_{i=1}^{N_1}\delta(\mathbf{x}-\mathbf{x}_i(t))$ and $\rho_2(\mathbf{x},t) = \sum_{i=1}^{N_2}\delta(\mathbf{x}-\mathbf{y}_i(t))$, the equations (1.7) evaluated for $\mathbf{v}_1$ at $\mathbf{x}_i$ and $\mathbf{v}_2$ at $\mathbf{y}_i$ reproduce (1.4) and (1.5) exactly, up to the time scaling introduced.

Before we proceed to the next section and linearize around the concentric ring steady state, it is worthwhile to consider the existence of such a steady state.

If $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$ parameterize concentric circles of radii R_1 and R_2, we can take $D = [-\pi, \pi)$ and $\mathbf{\Phi}_i(s) = \Theta(s) R_i \mathbf{e}_1$ (as in [87]) with

$$\Theta(s') = \begin{bmatrix} \cos s' & -\sin s' \\ \sin s' & \cos s' \end{bmatrix}.$$

With no motion in time, equations (1.9) and (1.10) give

$$\begin{aligned} 0 = \int_{-\pi}^{\pi} & g_1\left(\frac{1}{2}|\mathbf{\Phi}_1(s) - \mathbf{\Phi}_1(s')|^2\right)(\mathbf{\Phi}_1(s) - \mathbf{\Phi}_1(s'))f_1^0(s') + \\ & g_3\left(\frac{1}{2}|\mathbf{\Phi}_1(s) - \mathbf{\Phi}_2(s')|^2\right)(\mathbf{\Phi}_1(s) - \mathbf{\Phi}_2(s'))f_2^0(s')\,ds', \end{aligned}$$

$$\begin{aligned} 0 = \int_{-\pi}^{\pi} & g_2\left(\frac{1}{2}|\mathbf{\Phi}_2(s) - \mathbf{\Phi}_2(s')|^2\right)(\mathbf{\Phi}_2(s) - \mathbf{\Phi}_2(s'))f_2^0(s') + \\ & g_3\left(\frac{1}{2}|\mathbf{\Phi}_2(s) - \mathbf{\Phi}_1(s')|^2\right)(\mathbf{\Phi}_2(s) - \mathbf{\Phi}_1(s'))f_1^0(s')\,ds'. \end{aligned}$$

The (constant) densities and radii must satisfy

$$\int_{-\pi}^{\pi} f_i^0(s')\,ds' = \rho_i^s R_i = m_i,$$

where $m_i = \int_{\mathbb{R}^2} \rho_i$ is the total mass of species i. In the discrete case, $m_i = N_i$ and so $m_2/m_1 = N_2/N_1 = \mu$. Note that

$$f_i^0(s') = \rho_i^s \left|\frac{\partial \mathbf{\Phi}}{\partial s'}\right| = \rho_i^s R_i,$$

so $\mu = f_2^0/f_1^0$, and the above equations can be rewritten

$$\begin{aligned} 0 = \int_{-\pi}^{\pi} & g_1\left(\frac{1}{2}|\mathbf{\Phi}_1(s) - \mathbf{\Phi}_1(s')|^2\right)(\mathbf{\Phi}_1(s) - \mathbf{\Phi}_1(s')) + \\ & \mu g_3\left(\frac{1}{2}|\mathbf{\Phi}_1(s) - \mathbf{\Phi}_2(s')|^2\right)(\mathbf{\Phi}_1(s) - \mathbf{\Phi}_2(s'))\,ds', \end{aligned}$$

$$\begin{aligned} 0 = \int_{-\pi}^{\pi} & \mu g_2\left(\frac{1}{2}|\mathbf{\Phi}_2(s) - \mathbf{\Phi}_2(s')|^2\right)(\mathbf{\Phi}_2(s) - \mathbf{\Phi}_2(s')) + \\ & g_3\left(\frac{1}{2}|\mathbf{\Phi}_2(s) - \mathbf{\Phi}_1(s')|^2\right)(\mathbf{\Phi}_2(s) - \mathbf{\Phi}_1(s'))\,ds'. \end{aligned}$$

From the definition of $\mathbf{\Phi}_i$,

$$\mathbf{\Phi}_i(s) - \mathbf{\Phi}_j(s') = \Theta(s)[R_i I - R_j \Theta(s'-s)]\,\mathbf{e}_1 = \Theta(s)\begin{bmatrix} R_i - R_j\cos(s'-s) \\ -R_j\sin(s'-s) \end{bmatrix},$$

and so

$$|\mathbf{\Phi}_i(s) - \mathbf{\Phi}_j(s')|^2 = R_i^2 + R_j^2 - 2R_iR_j\cos(s'-s).$$

We may cancel $\Theta(s)$ from both equations and reparameterize the integrals so that s disappears as well, to reach

$$\begin{aligned} 0 = &\int_{-\pi}^{\pi} g_1(R_1^2(1-\cos s'))\begin{bmatrix} R_1 - R_1\cos s' \\ -R_1\sin s' \end{bmatrix} + \\ &\mu g_3\left(\frac{R_1^2+R_2^2}{2} - R_1R_2\cos s'\right)\begin{bmatrix} R_1 - R_2\cos s' \\ -R_2\sin s' \end{bmatrix} ds', \\ 0 = &\int_{-\pi}^{\pi} \mu g_2\big(R_2^2(1-\cos s')\big)\begin{bmatrix} R_2 - R_2\cos s' \\ -R_2\sin s' \end{bmatrix} + \\ &g_3\left(\frac{R_1^2+R_2^2}{2} - R_1R_2\cos s'\right)\begin{bmatrix} R_2 - R_1\cos s' \\ -R_1\sin s' \end{bmatrix} ds'. \end{aligned}$$

The second component of each integral cancels because it is odd on $(-\pi,\pi)$, and so we are left with

$$\begin{aligned} 0 = &\int_{-\pi}^{\pi} R_1 g_1(R_1^2(1-\cos s'))(1-\cos s') + \\ &\mu g_3\left(\frac{R_1^2+R_2^2}{2} - R_1R_2\cos s'\right)(R_1 - R_2\cos s')\,ds' \end{aligned} \tag{1.11a}$$

$$\begin{aligned} 0 = &\int_{-\pi}^{\pi} \mu R_2 g_2\big(R_2^2(1-\cos s')\big)(1-\cos s') + \\ &g_3\left(\frac{R_1^2+R_2^2}{2} - R_1R_2\cos s'\right)(R_2 - R_1\cos s')\,ds', \end{aligned} \tag{1.11b}$$

which determine R_1 and R_2. So long as (1.11a) and (1.11b) have solutions, the concentric ring steady state exists. Of course, the integrands of (1.11a) and (1.11b)

must be in $L^1[-\pi,\pi]$, which is true if $g_i(t)t^{1/2} \in L^1[0,1]$ and the g_i have no singularities away from the origin. Assuming also that $V_i(t)t^{-1/2} \in L^1[0,1]$, one can define

$$F(R_1,R_2) := \int_{-\pi}^{\pi} \frac{1}{2}V_1\big(R_1^2(1-\cos s')\big) + \frac{\mu^2}{2}V_2\big(R_2^2(1-\cos s')\big) + \mu V_3\left(\frac{R_1^2+R_2^2}{2} - R_1R_2\cos s'\right) ds'$$

and observe that (1.11a) arises as $\frac{\partial F}{\partial R_1} = 0$ and (1.11b) as $\frac{\partial F}{\partial R_2} = 0$. Then if g_1, g_2, and g_3 are continuous (except perhaps at the origin— for commonly encountered potentials such as the Morse or Lennard-Jones potentials, this is the case) (1.11a) and (1.11b) will be satisfied at a maximum or minimum of F. To show, then, that solutions R_1 and R_2 exist, it suffices to show that F attains a minimum for some $R_1, R_2 > 0$.

A general proof of this fact is difficult because the potentials V_i will vary, and in some cases a concentric ring solution may not exist. However, for all cases pursued below, (1.11a) and (1.11b) have solutions R_1, R_2 which do give rise to a steady state solution to (1.9) and (1.10).

1.5 Linearization & Eigenvalue Problem

Recall that the rings have been parameterized as $\mathbf{\Phi}_i(s) = \Theta(s)R_i\mathbf{e}_1$ (where Θ is a rotation matrix). Consider now a small perturbation of each ring in the form

$$\delta\mathbf{\Phi}_i(s) = \Theta_i(s)\big(R_i\mathbf{e}_1 + e^{\lambda t}\boldsymbol{\epsilon}_i(s)\big) = \mathbf{\Phi}_i(s) + \Theta(s)e^{\lambda t}\boldsymbol{\epsilon}_i(s)$$

so that (defining $\mathbf{A}_i, \mathbf{B}_i, \mathbf{C}, \mathbf{D}, \mathbf{E}, \mathbf{F}$)

$$\delta\mathbf{\Phi}_i(s) - \delta\mathbf{\Phi}_i(s') = (\mathbf{\Phi}_i(s) - \mathbf{\Phi}_i(s')) + e^{\lambda t}[\Theta(s)\boldsymbol{\epsilon}_i(s) - \Theta(s')\boldsymbol{\epsilon}_i(s')] =: \mathbf{A}_i + e^{\lambda t}\mathbf{B}_i,$$

$$\delta\mathbf{\Phi}_1(s) - \delta\mathbf{\Phi}_2(s') = (\mathbf{\Phi}_1(s) - \mathbf{\Phi}_2(s')) + e^{\lambda t}[\Theta(s)\boldsymbol{\epsilon}_1(s) - \Theta(s')\boldsymbol{\epsilon}_2(s')] =: \mathbf{C} + e^{\lambda t}\mathbf{D},$$

$$\delta\mathbf{\Phi}_2(s) - \delta\mathbf{\Phi}_1(s') = (\mathbf{\Phi}_2(s) - \mathbf{\Phi}_1(s')) + e^{\lambda t}[\Theta(s)\boldsymbol{\epsilon}_2(s) - \Theta(s')\boldsymbol{\epsilon}_1(s')] =: \mathbf{E} + e^{\lambda t}\mathbf{F}.$$

Linearizing (1.9) and (1.10), then canceling the factor of $e^{\lambda t}$ appearing in each term gives

$$\lambda\Theta(s)\boldsymbol{\epsilon}_1(s) = \int_0^{2\pi} g_1\left(\frac{1}{2}|\mathbf{A}_1|^2\right)\mathbf{B}_1 + \frac{dg_1}{ds}\left(\frac{1}{2}|\mathbf{A}_1|^2\right)(\mathbf{A}_1\cdot\mathbf{B}_1)\mathbf{A}_1 + \mu g_3\left(\frac{1}{2}|\mathbf{C}|^2\right)\mathbf{D} + \mu\frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{C}|^2\right)(\mathbf{C}\cdot\mathbf{D})\mathbf{C}\,ds',$$

$$\lambda\Theta(s)\boldsymbol{\epsilon}_2(s) = \int_0^{2\pi} \mu g_2\left(\frac{1}{2}|\mathbf{A}_2|^2\right)\mathbf{B}_2 + \mu\frac{dg_2}{ds}\left(\frac{1}{2}|\mathbf{A}_2|^2\right)(\mathbf{A}_2\cdot\mathbf{B}_2)\mathbf{A}_2 + g_3\left(\frac{1}{2}|\mathbf{E}|^2\right)\mathbf{F} + \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{E}|^2\right)(\mathbf{E}\cdot\mathbf{F})\mathbf{E}\,ds'.$$

To simplify calculations below, define $M = M(s,s') := \Theta^{-1}(s)\Theta(s')$, and

$$\begin{aligned}
\mathbf{u}_i &= \Theta^{-1}(s)\mathbf{A}_i & \tilde{\mathbf{u}}_i &= (I - M^T)R_i\mathbf{e}_1\\
\mathbf{v} &= \Theta^{-1}(s)\mathbf{C} & \tilde{\mathbf{v}} &= (R_2 I - R_1 M^T)\mathbf{e}_1\\
\mathbf{w} &= \Theta^{-1}(s)\mathbf{E} & \tilde{\mathbf{w}} &= (R_1 I - R_2 M)\mathbf{e}_1.
\end{aligned}$$

Then

$$\begin{aligned}
\Theta^{-1}(s)\mathbf{A}_i &= (I - M)\,R_i\mathbf{e}_1 = \mathbf{u}_i & \Theta^{-1}(s)\mathbf{B}_i &= \boldsymbol{\epsilon}_i(s) - M\boldsymbol{\epsilon}_i(s')\\
\Theta^{-1}(s)\mathbf{C} &= (R_1 I - R_2 M)\mathbf{e}_1 = \mathbf{v} & \Theta^{-1}(s)\mathbf{D} &= \boldsymbol{\epsilon}_1(s) - M\boldsymbol{\epsilon}_2(s')\\
\Theta^{-1}(s)\mathbf{E} &= (R_2 I - R_1 M)\mathbf{e}_1 = \mathbf{w} & \Theta^{-1}(s)\mathbf{F} &= \boldsymbol{\epsilon}_2(s) - M\boldsymbol{\epsilon}_1(s')
\end{aligned}$$

and because Θ and M are unitary,

$$\begin{aligned}
\mathbf{A}_i\cdot\mathbf{B}_i &= [\Theta(s)(I-M)R_i\mathbf{e}_1]\cdot[\Theta(s)\boldsymbol{\epsilon}_i(s)] - [\Theta(s)(I-M)R_i\mathbf{e}_1]\cdot[\Theta(s)M\boldsymbol{\epsilon}_i(s')]\\
&= (I-M)R_i\mathbf{e}_1\cdot\boldsymbol{\epsilon}_i(s) - (I-M)R_i\mathbf{e}_1\cdot M\boldsymbol{\epsilon}_i(s')\\
&= (I-M)R_i\mathbf{e}_1\cdot\boldsymbol{\epsilon}_i(s) + (I-M^T)R_i\mathbf{e_1}\cdot\boldsymbol{\epsilon}_i(s'),\\
&= \mathbf{u}_i\cdot\boldsymbol{\epsilon}_i(s) + \tilde{\mathbf{u}}_i\cdot\boldsymbol{\epsilon}_i(s'),\\
\mathbf{C}\cdot\mathbf{D} &= (R_1 I - R_2 M)\mathbf{e}_1\cdot\boldsymbol{\epsilon}_1(s) + (R_2 I - R_1 M^T)\mathbf{e}_1\cdot\boldsymbol{\epsilon}_2(s')\\
&= \mathbf{v}\cdot\boldsymbol{\epsilon}_1(s) + \tilde{\mathbf{v}}\cdot\boldsymbol{\epsilon}_2(s'),\\
\mathbf{E}\cdot\mathbf{F} &= (R_2 I - R_1 M)\mathbf{e}_1\cdot\boldsymbol{\epsilon}_2(s) + (R_1 I - R_2 M^T)\mathbf{e}_1\cdot\boldsymbol{\epsilon}_1(s')\\
&= \mathbf{w}\cdot\boldsymbol{\epsilon}_2(s) + \tilde{\mathbf{w}}\cdot\boldsymbol{\epsilon}_1(s').
\end{aligned}$$

Finally,

$$|\mathbf{A}_i| = |\Theta^{-1}\mathbf{A}_i| = |\mathbf{u}_i|$$
$$|\mathbf{C}_i| = |\Theta^{-1}\mathbf{C}_i| = |\mathbf{v}_i|$$
$$|\mathbf{E}_i| = |\Theta^{-1}\mathbf{E}_i| = |\mathbf{w}_i|.$$

In terms of these quantities, multiplying the linearized equations by Θ^{-1} and collecting terms multiplied by $\boldsymbol{\epsilon}_i(s)$ and $\boldsymbol{\epsilon}_i(s')$ leaves

$$\begin{aligned}\lambda\boldsymbol{\epsilon}_1(s) = \int_{-\pi}^{\pi} \Bigg[& g_1\left(\frac{1}{2}|\mathbf{u}_1|^2\right) I + \frac{dg_1}{ds}\left(\frac{1}{2}|\mathbf{u}_1|^2\right) \mathbf{u}_1 \otimes \mathbf{u}_1 + \\ & \mu g_3\left(\frac{1}{2}|\mathbf{v}|^2\right) I + \mu\frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{v}|^2\right) \mathbf{v} \otimes \mathbf{v}\Bigg] \boldsymbol{\epsilon}_1(s)\, ds' \\ + \int_{-\pi}^{\pi} & \left[-g_1\left(\frac{1}{2}|\mathbf{u}_1|^2\right) M + \frac{dg_1}{ds}\left(\frac{1}{2}|\mathbf{u}_1|^2\right) \mathbf{u}_1 \otimes \tilde{\mathbf{u}}_1\right] \boldsymbol{\epsilon}_1(s')\, ds' \\ + \int_{-\pi}^{\pi} & \left[-\mu g_3\left(\frac{1}{2}|\mathbf{v}|^2\right) M + \mu\frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{v}|^2\right) \mathbf{v} \otimes \tilde{\mathbf{v}}\right] \boldsymbol{\epsilon}_2(s')\, ds',\end{aligned} \tag{1.12}$$

$$\begin{aligned}\lambda\boldsymbol{\epsilon}_2(s) = \int_{-\pi}^{\pi} \Bigg[& \mu g_2\left(\frac{1}{2}|\mathbf{u}_2|^2\right) I + \mu\frac{dg_2}{ds}\left(\frac{1}{2}|\mathbf{u}_2|^2\right) \mathbf{u}_2 \otimes \mathbf{u}_2 + \\ & g_3\left(\frac{1}{2}|\mathbf{w}|^2\right) I + \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{w}|^2\right) \mathbf{w} \otimes \mathbf{w}\Bigg] \boldsymbol{\epsilon}_2(s)\, ds' \\ + \int_{-\pi}^{\pi} & \left[-\mu g_2\left(\frac{1}{2}|\mathbf{u}_2|^2\right) M + \mu\frac{dg_2}{ds}\left(\frac{1}{2}|\mathbf{u}_2|^2\right) \mathbf{u}_2 \otimes \tilde{\mathbf{u}}_2\right] \boldsymbol{\epsilon}_2(s')\, ds' \\ + \int_{-\pi}^{\pi} & \left[-g_3\left(\frac{1}{2}|\mathbf{w}|^2\right) M + \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{w}|^2\right) \mathbf{w} \otimes \tilde{\mathbf{w}}\right] \boldsymbol{\epsilon}_1(s')\, ds'.\end{aligned} \tag{1.13}$$

Explicitly,

$$\frac{1}{2}|\mathbf{u}_i(s-s')|^2 = R_i^2(1-\cos(s-s'))$$
$$\frac{1}{2}|\mathbf{v}(s-s')|^2 = \frac{1}{2}|\mathbf{w}(s-s')|^2 = \frac{1}{2}\left[R_1^2 + R_2^2 - 2R_1R_2\cos(s-s')\right]$$

$$\mathbf{u}_i = R_i\begin{bmatrix}1-\cos(s'-s)\\ -\sin(s'-s)\end{bmatrix}, \quad \tilde{\mathbf{u}}_i = R_i\begin{bmatrix}1-\cos(s'-s)\\ \sin(s'-s)\end{bmatrix}$$
$$\mathbf{v} = \begin{bmatrix}R_1 - R_2\cos(s'-s)\\ -R_2\sin(s'-s)\end{bmatrix}, \quad \tilde{\mathbf{v}} = \begin{bmatrix}R_1 - R_2\cos(s'-s)\\ R_2\sin(s'-s)\end{bmatrix}$$

$$\mathbf{w} = \begin{bmatrix} R_2 - R_1\cos(s'-s) \\ -R_1\sin(s'-s) \end{bmatrix}, \quad \tilde{\mathbf{w}} = \begin{bmatrix} R_2 - R_1\cos(s'-s) \\ R_1\sin(s'-s) \end{bmatrix}$$

and $u \otimes v$ denotes the matrix with i, j entry $u_i v_j$.

Note that all the above matrices have even, periodic entries along the diagonals and odd, periodic entries off. With this in mind, consider (1.12) rewritten as

$$\lambda \boldsymbol{\epsilon}_1(s) = \int_0^{2\pi} M^1(s'-s)\, ds' \boldsymbol{\epsilon}_1(s) + \int_0^{2\pi} M^2(s'-s)\boldsymbol{\epsilon}_1(s')\, ds' + \int_0^{2\pi} M^3(s'-s)\boldsymbol{\epsilon}_2(s')\, ds'$$

(the superscripts are used to distinguish matrices, not as powers) where the diagonal entries of M^1, M^2, and M^3 are even and periodic, and off-diagonal entries are odd and periodic. It follows that

$$\int_0^{2\pi} M^1(s'-s)\, ds' \boldsymbol{\epsilon}_1(s)$$

is a constant diagonal matrix times $\boldsymbol{\epsilon}_1(s)$. For the other two terms, we hope for similar results to yield an eigenvalue problem in $\boldsymbol{\epsilon}_i$ and λ. Using the ansatz

$$\boldsymbol{\epsilon}_1(s) = \begin{bmatrix} x_1\cos(ns) \\ x_2\sin(ns) \end{bmatrix}, \quad \boldsymbol{\epsilon}_2(s) = \begin{bmatrix} y_1\cos(ns) \\ y_2\sin(ns) \end{bmatrix}$$

similar to that of [51], we compute the terms above involving M^2 and M^3:

$$\int_0^{2\pi} M^2(s-s')\boldsymbol{\epsilon}_1(s')\, ds' = \begin{bmatrix} \int_0^{2\pi} M^2_{11}(s-s')x_1\cos(ns') + M^2_{12}(s-s')x_2\sin(ns')\, ds' \\ \int_0^{2\pi} M^2_{21}(s-s')x_1\cos(ns') + M^2_{22}(s-s')x_2\sin(ns')\, ds' \end{bmatrix}.$$

The first entry is a linear combination of

$$\begin{aligned} \int_0^{2\pi} M^2_{11}(s'-s)\cos(ns')\, ds' &= \int_0^{2\pi} M^2_{11}(\theta)\cos(n\theta + ns)\, d\theta \\ &= \cos(ns)\int_0^{2\pi} M^2_{11}(\theta)\cos(n\theta)\, d\theta - \sin(ns)\int_0^{2\pi} M^2_{11}(\theta)\sin(n\theta)\, d\theta \\ &= \cos(ns)\int_0^{2\pi} M^2_{11}(\theta)\cos(n\theta)\, d\theta + 0 \qquad \text{(because } M^2_{11} \text{ is even)} \\ &\propto \cos(ns), \end{aligned} \tag{1.14a}$$

and

$$\begin{aligned}\int_0^{2\pi} M_{12}^2(s'-s)\sin(ns')\,ds' &= \int_0^{2\pi} M_{12}^2(\theta)\sin(n\theta+ns)\,d\theta \\ &= \cos(ns)\int_0^{2\pi} M_{12}^2(\theta)\sin(n\theta)\,d\theta + \sin(ns)\int_0^{2\pi} M_{12}^2(\theta)\cos(n\theta)\,d\theta \\ &= \cos(ns)\int_0^{2\pi} M_{12}^2(\theta)\sin(n\theta)\,d\theta + 0 \qquad \text{(because } M_{12}^2 \text{ is odd)} \\ &\propto \cos(ns); \end{aligned} \tag{1.14b}$$

the second entry,

$$\int_0^{2\pi} M_{21}^2(s'-s) \propto \sin(ns) \tag{1.14c}$$

and

$$\int_0^{2\pi} M_{22}^2(s'-s) \propto \sin(ns) \tag{1.14d}$$

after similar calculations. All together,

$$\int_0^{2\pi} M^2(s-s')\boldsymbol{\epsilon}_1(s')\,ds' = \mathbf{a}(n)\boldsymbol{\epsilon}_1(s)$$

where $\mathbf{a}$ is a diagonal matrix. The third term will be similar and will give a diagonal matrix multiple of $\boldsymbol{\epsilon}_2$, so that the equation for $\boldsymbol{\epsilon}_1$ becomes

$$\lambda \begin{bmatrix} x_1\cos(ns) \\ x_2\sin(ns) \end{bmatrix} = \mathbf{a}(n) \begin{bmatrix} x_1\cos(ns) \\ x_2\sin(ns) \end{bmatrix} + \mathbf{b}(n) \begin{bmatrix} y_1\cos(ns) \\ y_2\sin(ns) \end{bmatrix}$$

($\mathbf{a}$ and $\mathbf{b}$ are matrix-valued functions) and the equation for $\boldsymbol{\epsilon}_2$ is

$$\lambda \begin{bmatrix} y_1\cos(ns) \\ y_2\sin(ns) \end{bmatrix} = \mathbf{c}(n) \begin{bmatrix} y_1\cos(ns) \\ y_2\sin(ns) \end{bmatrix} + \mathbf{d}(n) \begin{bmatrix} x_1\cos(ns) \\ x_2\sin(ns) \end{bmatrix}.$$

Comparing coefficients of $\cos(ns)$ and $\sin(ns)$ results in an eigenvalue problem for x_1, x_2, y_1, and y_2:

$$\lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \left(\mathbf{M}^1 + \mathbf{M}^2(n)\right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \mathbf{M}^3(n) \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$
$$\lambda \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{M}^4(n) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \left(\mathbf{M}^5 + \mathbf{M}^6(n)\right) \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$

or

$$\mathbf{E}(n) \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \lambda(n) \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}, \quad \text{with } \mathbf{E} = \begin{bmatrix} \mathbf{M}^1 + \mathbf{M}^2 & \mathbf{M}^3 \\ \mathbf{M}^4 & \mathbf{M}^5 + \mathbf{M}^6 \end{bmatrix}. \tag{1.15}$$

$\mathbf{M}^1, \ldots, \mathbf{M}^6$ are computed as in (1.14a—1.14d) and are shown below; here, $\mathbf{u}_i$, $\mathbf{v}$, and $\mathbf{w}$ are functions of θ. $\mathbf{M}^1$ and $\mathbf{M}^5$ are diagonal and do not depend on n.

The first two matrices determine the stability of the species I particle ring with respect to frequency n perturbations, with the species II ring remaining fixed:

$$\begin{aligned}
\mathbf{M}^1_{11} &= \int_{-\pi}^{\pi} g_1\left(\frac{1}{2}|\mathbf{u}_1|^2\right) + \frac{dg_1}{ds}\left(\frac{1}{2}|\mathbf{u}_1|^2\right)\left(R_1 - R_1\cos(s')\right)^2 + \\
&\qquad \mu g_3\left(\frac{1}{2}|\mathbf{v}|^2\right) + \mu\frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{v}|^2\right)\left(R_1 - R_2\cos(s')\right)^2 ds' \\
\mathbf{M}^1_{22} &= \int_{-\pi}^{\pi} g_1\left(\frac{1}{2}|\mathbf{u}_1|^2\right) + \frac{dg_1}{ds}\left(\frac{1}{2}|\mathbf{u}_1|^2\right) R_1^2\sin^2(s') + \\
&\qquad \mu g_3\left(\frac{1}{2}|\mathbf{v}|^2\right) + \mu\frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{v}|^2\right) R_2^2\sin^2(s')\, ds'.
\end{aligned}$$

$$\begin{aligned}
\mathbf{M}^2_{11}(n) &= \int_{-\pi}^{\pi}\left[-g_1\left(\frac{1}{2}|\mathbf{u}_1|^2\right)\cos(\theta) + R_1^2\frac{dg_1}{ds}\left(\frac{1}{2}|\mathbf{u}_1|^2\right)\left(1-\cos(\theta)\right)^2\right]\cos(n\theta)\, d\theta \\
\mathbf{M}^2_{12}(n) &= \int_{-\pi}^{\pi}\left[g_1\left(\frac{1}{2}|\mathbf{u}_1|^2\right)\sin(\theta) + R_1^2\frac{dg_1}{ds}\left(\frac{1}{2}|\mathbf{u}_1|^2\right)\left(1-\cos(\theta)\right)\sin(\theta)\right]\sin(n\theta)\, d\theta \\
\mathbf{M}^2_{21}(n) &= \mathbf{M}^2_{12} \\
\mathbf{M}^2_{22}(n) &= \int_{-\pi}^{\pi}\left[-g_1\left(\frac{1}{2}|\mathbf{u}_1|^2\right)\cos(\theta) - R_1^2\frac{dg_1}{ds}\left(\frac{1}{2}|\mathbf{u}_1|^2\right)\sin^2(\theta)\right]\cos(n\theta)\, d\theta \\
&= 0
\end{aligned}$$

after integrating by parts and using (1.11a).

Due to symmetry in the problem, the off-diagonal blocks $\mathbf{M}^3$ and $\mathbf{M}^4$ are similar. $\mathbf{M}^3$ represents the effect of a perturbation of the species II particles on the ring of species I particles; $\mathbf{M}^4$, the effect of a perturbation of the species I ring on the species II ring:

$$\mathbf{M}^3_{11}(n) = \mu \int_{-\pi}^{\pi} \left[-g_3\left(\frac{1}{2}|\mathbf{v}|^2\right)\cos(\theta) + \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{v}|^2\right)(R_1 - R_2\cos(\theta))^2 \right] \cos(n\theta)\, d\theta$$

$$\mathbf{M}^3_{12}(n) = \mu \int_{-\pi}^{\pi} \left[g_3\left(\frac{1}{2}|\mathbf{v}|^2\right)\sin(\theta) + \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{v}|^2\right) R_2(R_1 - R_2\cos(\theta))\sin(\theta) \right] \sin(n\theta) d\theta$$

$$\mathbf{M}^3_{21}(n) = \mathbf{M}^3_{12}$$

$$\mathbf{M}^3_{22}(n) = \mu \int_{-\pi}^{\pi} \left[-g_3\left(\frac{1}{2}|\mathbf{v}|^2\right)\cos(\theta) - \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{v}|^2\right) R_2^2\sin^2(\theta) \right] \cos(n\theta)\, d\theta.$$

$$\mathbf{M}^4_{11}(n) = \int_{-\pi}^{\pi} \left[-g_3\left(\frac{1}{2}|\mathbf{w}|^2\right)\cos(\theta) + \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{w}|^2\right)(R_2 - R_1\cos(\theta))^2 \right] \cos(n\theta)\, d\theta$$

$$\mathbf{M}^4_{12}(n) = \int_{-\pi}^{\pi} \left[g_3\left(\frac{1}{2}|\mathbf{w}|^2\right)\sin(\theta) + \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{w}|^2\right) R_1(R_2 - R_1\cos(\theta))\sin(\theta) \right] \sin(n\theta) d\theta$$

$$\mathbf{M}^4_{21}(n) = \mathbf{M}^4_{12}$$

$$\mathbf{M}^4_{22}(n) = \int_{-\pi}^{\pi} \left[-g_3\left(\frac{1}{2}|\mathbf{w}|^2\right)\cos(\theta) - \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{w}|^2\right) R_1^2\sin^2(n\theta) \right] \cos(n\theta)\, d\theta.$$

The final two matrices $\mathbf{M}^5$ and $\mathbf{M}^6$ are analogous to $\mathbf{M}^1$ and $\mathbf{M}^2$, except they determine the stability of the species II ring:

$$\begin{aligned}
\mathbf{M}^5_{11} = \int_{-\pi}^{\pi} & \mu g_2\left(\frac{1}{2}|\mathbf{u}_2|^2\right) + \mu\frac{dg_2}{ds}\left(\frac{1}{2}|\mathbf{u}_2|^2\right)(R_2 - R_2\cos(s'))^2 + \\
& g_3\left(\frac{1}{2}|\mathbf{w}|^2\right) + \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{w}|^2\right)(R_2 - R_1\cos(s'))^2\, ds' \\
\mathbf{M}^5_{22} = \int_{-\pi}^{\pi} & \mu g_2\left(\frac{1}{2}|\mathbf{u}_2|^2\right) + \mu\frac{dg_2}{ds}\left(\frac{1}{2}|\mathbf{u}_2|^2\right) R_2^2\sin^2(s') + \\
& g_3\left(\frac{1}{2}|\mathbf{w}|^2\right) + \frac{dg_3}{ds}\left(\frac{1}{2}|\mathbf{w}|^2\right) R_1^2\sin^2(s')\, ds' = 0
\end{aligned}$$

after integrating by parts and using (1.11b), and

$$\mathbf{M}^6_{11}(n) = \mu \int_{-\pi}^{\pi} \left[-g_2\left(\frac{1}{2}|\mathbf{u}_2|^2\right)\cos(\theta) + R_2^2 \frac{dg_2}{ds}\left(\frac{1}{2}|\mathbf{u}_2|^2\right)(1-\cos(\theta))^2 \right] \cos(n\theta)\, d\theta$$

$$\mathbf{M}^6_{12}(n) = \mu \int_{-\pi}^{\pi} \left[g_2\left(\frac{1}{2}|\mathbf{u}_2|^2\right)\sin(\theta) + R_2^2 \frac{dg_2}{ds}\left(\frac{1}{2}|\mathbf{u}_2|^2\right)(1-\cos(\theta))\sin(\theta) \right] \sin(n\theta) d\theta$$

$$\mathbf{M}^6_{21}(n) = \mathbf{M}^6_{21}$$

$$\mathbf{M}^6_{22}(n) = \mu \int_{-\pi}^{\pi} \left[-g_2\left(\frac{1}{2}|\mathbf{u}_2|^2\right)\cos(\theta) - R_2^2 \frac{dg_2}{ds}\left(\frac{1}{2}|\mathbf{u}_2|^2\right)\sin^2(\theta) \right] \cos(n\theta)\, d\theta.$$

1.6 Linear Well-posedness

Here we consider the limit of the eigenvalue problem (1.15) as $n \to \infty$. The goal is *linear well-posedness*; that is, to determine when the eigenvalues $\lambda(n)$ of (1.15) satisfy $\lambda(n) < 0$ as $n \to \infty$. That $\lambda(n) \to 0$ as $n \to \infty$ follows immediately from the Riemann-Lebesgue lemma; the requirement that the eigenvalues approach zero from below is important because it demonstrates that all but finitely many modes are stable. Intuitively, if modes of arbitrarily high frequency are unstable, the co-dimension one curve will break apart and the density will form a fully two-dimensional pattern.

Theorem 1.6.1 (Linear well-posedness)**.** *Assume that the forces have power series representations*

$$g_1(s) = a_0 s^{p_0} + a_1 s^{p_1} + \dots \tag{1.16a}$$

$$g_2(s) = b_0 s^{q_0} + b_1 s^{q_1} + \dots \tag{1.16b}$$

$$g_3(s) = c_0 s^{r_0} + c_1 s^{r_1} + \dots \tag{1.16c}$$

with $a_0, b_0 > 0$ and $c_0 \neq 0$, valid in some neighborhood of the origin, where $p_0 < p_1 < \dots$ etc. Define $\alpha = \mathbf{M}^1_{11}$ and $\beta = \mathbf{M}^5_{11}$.

If $R_1 \neq R_2$, then the concentric ring solution to (1.9, 1.10) is linearly well-

posed if and only if

$$\alpha < 0, \quad p_0 \in \left(-\frac{1}{2}, 0\right) \cup \bigcup_{n=0}^{\infty} (2n+1, 2n+2), \tag{1.17a}$$

$$\beta < 0, \quad q_0 \in \left(-\frac{1}{2}, 0\right) \cup \bigcup_{n=0}^{\infty} (2n+1, 2n+2). \tag{1.17b}$$

If $R_1 = R_2$, $g_1 = g_2$, and $\mu = 1$, the concentric ring solution to (1.9, 1.10) is linearly well-posed if and only if (1.17a) holds and either r_0 is a nonnegative integer or $r_0 > p_0$.

Before moving on to the proof, a remark is in order. Theorem 1.6.1 as stated does not cover the cases when $R_1 = R_2$ but $g_1 \neq g_2$ or $\mu \neq 1$. However, (1.11a) and (1.11b) point out that unless $\mu = 1$ and $g_1 = g_2$, it is very unlikely that $R_1 = R_2$; for two arbitrary potentials g_1, g_2 and ratio μ, it is a measure-zero type event that the radii equations would have such solutions.

Proof of Theorem 1.6.1. The analysis relies primarily on asymptotic expressions for the integrals occurring in $\mathbf{M}^1, \ldots, \mathbf{M}^6$, which necessitates the assumptions of (1.16). Substituting (1.16) into the formulas for $\mathbf{M}^1, \ldots, \mathbf{M}^6$ leaves an eigenvalue problem where each entry of $\mathbf{E}$ is a potentially infinite sum of integrals. However, showing that $\mathbf{E}$ has negative eigenvalues is equivalent to showing its leading minors alternate sign, and it is easy to see that in each entry of $\mathbf{E}$, only those terms which decay most slowly will affect the eigenvalues in the limit.

In practice, the values $\alpha = \mathbf{M}^1_{11}$ and $\beta = \mathbf{M}^2_{11}$ must be evaluated analytically or numerically, because they are independent of n. Other entries of $\mathbf{E}$ may be evaluated asymptotically, and considering one of these entries gives an idea of

how to proceed:

$$\begin{aligned}\mathbf{M}_{11}^2 &\sim a_0 R_1^{2p_0} \int_{-\pi}^{\pi} -(1-\cos\theta)^{p_0}\cos(\theta)\cos(n\theta) + p_0(1-\cos\theta)^{p_0+1}\cos(n\theta)\,d\theta \\ &= a_0 R_1^{2p_0} \int_{-\pi}^{\pi} -(1-\cos\theta)^{p_0}\frac{1}{2}[\cos(n-1)\theta + \cos(n+1)\theta] + \\ &\quad p_0(1-\cos\theta)^{p_0+1}\cos(n\theta)\,d\theta\end{aligned}$$

where we used the trig identity

$$\cos(x)\cos(y) = \frac{\cos(x-y)+\cos(x+y)}{2}.$$

By a similar use similar identities, it turns out that all entries of $\mathbf{E}$ reduce to linear combinations of one integral:

$$I(c,n,p) = \int_{-\pi}^{\pi} (c-\cos\theta)^p \cos n\theta\,d\theta,$$

and we are interested in the behavior of I for fixed c and p as $n \to \infty$.

For $c = 1$ and $p > -1/2$ (which is necessary for the integrals to converge), an explicit formula with asymptotics is available from [87]:

$$I(1,n,p) \sim \frac{-C(p)\sin(\pi p)}{n^{2p+1}} \tag{1.18}$$

where $C(p) > 0$ is a positive constant depending on p. This asymptotic form may also be arrived at by stationary phase analysis.

For $c > 1$, it can be shown readily via integration by parts that I decays faster than any polynomial: for any integer k, there exists a constant $C(c,p,k)$ such that

$$|I(c,n,p)| < C(c,p,k)n^{-k}. \tag{1.19}$$

For $\mathbf{M}^2$ and $\mathbf{M}^6$, (1.18) gives the relevant rates of decay. $\mathbf{M}^3$ and $\mathbf{M}^4$ are more complicated and the analysis breaks into cases.

Case I: $R_1 \neq R_2$. In this case (1.19) shows that $\mathbf{M}^3$ and $\mathbf{M}^4$ approach zero faster than any of the other entries of $\mathbf{E}$, and (1.15) asymptotically decouples into two quasi-single species problems

$$(\mathbf{M}^1 + \mathbf{M}^2)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda(n)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad (\mathbf{M}^5 + \mathbf{M}^6)\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \lambda(n)\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \tag{1.20}$$

Before moving on, it is worth pointing out that the decoupling has a pleasant physical interpretation. $\mathbf{M}^3$ represents the effects of perturbations of the species II particle ring on the species I particle ring, and vice versa for $\mathbf{M}^4$. The fact that these vanish from (1.15) as $n \to \infty$ means that, when $R_1 \neq R_2$, high frequency perturbations of the ring of species I particles have no effect on the eventual stability of the ring of species II (and vice versa).

For each of the problems in (1.20) to have negative eigenvalues, it is necessary and sufficient that

$$\begin{aligned} &\operatorname{tr}(\mathbf{M}^1 + \mathbf{M}^2) < 0, \quad \det(\mathbf{M}^1 + \mathbf{M}^2) > 0, \\ &\operatorname{tr}(\mathbf{M}^5 + \mathbf{M}^6) < 0, \quad \det(\mathbf{M}^5 + \mathbf{M}^6) > 0. \end{aligned} \tag{1.21}$$

We turn first to $\mathbf{M}^2$: substituting in the assumed power series representation of g_1 and discarding all but the most slowly decaying terms,

$$\begin{aligned} \mathbf{M}^2_{11} &\sim a_0 R_1^{2p_0} \int_{-\pi}^{\pi} -(1-\cos\theta)^{p_0}\cos(\theta)\cos(n\theta) + p_0(1-\cos\theta)^{p_0+1}\cos(n\theta)\,d\theta \\ &= a_0 R_1^{2p_0} \int_{-\pi}^{\pi} -(1-\cos\theta)^{p_0}\frac{1}{2}[\cos(n-1)\theta + \cos(n+1)\theta] + \\ &\qquad p_0(1-\cos\theta)^{p_0+1}\cos(n\theta)\,d\theta \\ &= a_0 R_1^{2p_0}\left[-\frac{1}{2}I(1,n-1,p_0) + p_0 I(1,n,p_0+1) - \frac{1}{2}I(1,n+1,p_0)\right] \\ &\sim \frac{a_0 R_1^{2p_0} C(p_0)\sin(\pi p_0)}{2}\left[\frac{1}{(n-1)^{2p_0+1}} + \frac{1}{(n+1)^{2p_0+1}}\right] \\ &\sim \frac{a_0 R_1^{2p_0} C(p_0)\sin(\pi p_0)}{n^{2p_0+1}}, \end{aligned}$$

$$
\begin{aligned}
\mathbf{M}^2_{12} &\sim a_0 R_1^{2p_0} \int_{-\pi}^{\pi} (1-\cos\theta)^{p_0} \sin(\theta)\sin(n\theta) + p_0(1-\cos\theta)^{p_0}\sin(\theta)\sin(n\theta)\,d\theta \\
&= a_0 R_1^{2p_0}\left[\frac{p_0+1}{2} I(1,n-1,p_0) - \frac{p_0+1}{2} I(1,n+1,p_0)\right] \\
&= -\frac{a_0 R_1^{2p_0} C(p_0)\sin(\pi p_0)(p_0+1)}{2}\left[\frac{1}{(n-1)^{2p_0+1}} - \frac{1}{(n+1)^{2p_0+1}}\right] \\
&\sim -\frac{a_0 R_1^{2p_0} C(p_0)\sin(\pi p_0)(p_0+1)(2p_0+1)}{n^{2p_0+2}},
\end{aligned}
$$

and $\mathbf{M}^2_{21} = \mathbf{M}^2_{12}$. The final entry is treated by integration by parts:

$$
\begin{aligned}
\mathbf{M}^2_{22} &= \int_{-\pi}^{\pi} -g_1(R_1^2(1-\cos\theta))\cos\theta\cos(n\theta) \\
&\qquad - \frac{d}{d\theta}\big[g_1(R_1^2(1-\cos\theta))\big]\sin\theta\cos(n\theta)\,d\theta \\
&= -n\int_{-\pi}^{\pi} g_1(R_1^2(1-\cos\theta))\sin\theta\sin(n\theta)\,d\theta \\
&\sim -\frac{na_0 R_1^{2p_0}}{2}\int_{-\pi}^{\pi}(1-\cos\theta)^{p_0}[\cos(n-1)\theta - \cos(n+1)\theta] \\
&= -\frac{na_0 R_1^{2p_0}}{2}[I(1,n-1,p_0) - I(1,n+1,p_0)] \\
&= \frac{na_0 R_1^{2p_0} C(p_0)\sin(\pi p_0)}{2}\left[\frac{1}{(n-1)^{2p_0+1}} - \frac{1}{(n+1)^{2p_0+1}}\right] \\
&\sim \frac{a_0 R_1^{2p_0} C(p_0)\sin(\pi p_0)(2p_0+1)}{n^{2p_0+1}}.
\end{aligned}
$$

The analogous work for $\mathbf{M}^6$ looks almost exactly the same, except with q_0 replacing p_0.

With those expansions in hand, one can asymptotically compute the terms appearing in (1.21):

$$
\begin{aligned}
\operatorname{tr}(\mathbf{M}^1 + \mathbf{M}^2) &\sim \alpha \\
\det(\mathbf{M}^1 + \mathbf{M}^2) &\sim \alpha \mathbf{M}^2_{22}
\end{aligned}
$$

and so we need only require that $\alpha < 0$ and $\mathbf{M}^2_{22} < 0$. The asymptotic expression for the latter is negative so long as $\sin(\pi p_0)$ is, and this leads to (1.17a). The problem for $\mathbf{M}^5 + \mathbf{M}^6$ is nearly identical, and yields (1.17b) in exactly the same

way. It is worth noting here that the criteria for linear well-posedness are very similar to those for the single-species case explored in [51] and [87].

Case II: $R_1 = R_2 =: R$. As mentioned earlier, this is very unlikely unless $g_1 = g_2$ and $\mu = 1$; so, we will assume that is the case. Then $\mathbf{M}^1 = \mathbf{M}^5$, $\mathbf{M}^2 = \mathbf{M}^6$, and $\mathbf{M}^3 = \mathbf{M}^4$, so $\mathbf{E}$ simplifies; however, the rate of decay of $\mathbf{M}^3$ is not as fast now and so it must be taken into account. We determine when $\mathbf{E}$ has negative eigenvalues by checking when its leading minors alternate sign. Asymptotics for $\mathbf{M}^3$ are necessary, but $\mathbf{M}^3$ has the same form as $\mathbf{M}^2$ with g_3 replacing g_1 and R replacing R_1:

$$\mathbf{M}^3_{11} \sim \frac{c_0 R^{2r_0} C(r_0) \sin(\pi r_0)}{n^{2r_0+1}},$$

$$\mathbf{M}^3_{12} = \mathbf{M}^3_{21} \sim -\frac{c_0 R^{2r_0} C(r_0) \sin(\pi r_0)(r_0+1)(2r_0+1)}{n^{2r_0+2}},$$

$$\mathbf{M}^3_{22} \sim \frac{c_0 R^{2r_0} C(r_0) \sin(\pi r_0)(2r_0+1)}{n^{2r_0+1}}.$$

The first minor of $\mathbf{E}$ is then $(\mathbf{M^1} + \mathbf{M}^2(n))_{11} \to \alpha$ as $n \to \infty$, so we must require that $\alpha < 0$.

The second minor is $\det(\mathbf{M^1} + \mathbf{M}^2(n)) \sim \alpha \mathbf{M}^2_{22}$ (see case I), so we require that (1.17a) holds.

The third minor begins to include terms from the cross-particle interaction force g_3, and works out to be (to leading order in n)

$$\alpha^2 \mathbf{M}^2_{22} - \alpha(\mathbf{M}^3_{12})^2 = -C_1 n^{-(2p_0+1)} + C_2 \sin^2(\pi r_0) n^{-(4r_0+4)}$$

where C_1 and C_2 are some positive constants with respect to n. We have already required for the first and second minors that $\alpha < 0$ and $\mathbf{M}^2_{22} < 0$, which is why the first term is negative and the second is positive. So we must require that

$$r_0 \text{ is a nonnegative integer} \quad \text{or} \quad 2p_0 + 1 < 4r_0 + 4.$$

The fourth minor works out to be (again to leading order)

$$\alpha^2\left[(\mathbf{M}_{22}^2)^2 - (\mathbf{M}_{22}^3)^2\right] = C_1 n^{-2(2p_0+1)} - C_2 \sin^2(\pi r_0) n^{-2(2r_0+1)}$$

where C_1 and C_2 again denote positive constants. So we must require that

$$r_0 \text{ is a nonnegative integer} \quad \text{or} \quad r_0 > p_0.$$

These restrictions also imply that the third minor is negative. This last minor gives the final requirement for the double ring solution to be linearly well-posed and yields the criterion in the theorem for the $R_1 = R_2$ case. □

1.7 Numerical Examples

All numerical solutions here and in figure 1.1 were computed using a simple forward Euler scheme with an adaptive time step chosen as large as possible while requiring that the energy of the system (1.3) decreases at each time step. A scheme with higher order accuracy is not necessary, since we only seek a minimizer of the energy (1.3). Alternatively, choosing a time step based on an estimate of the local truncation error (as in [52]) is also efficient and yields the same results. All initial conditions are taken to be independently, uniformly distributed on a square.

1.7.1 Mixing or Separation of the Two Species

The theoretical predictions agree very well with numerical observations. Figure 1.4 shows an example in which the two particle species may mix or segregate based on the relative strengths of the inter-species and intra-species interactions. The numerical destabilization of the alternating ring structure and appearance of mode two instability coincides exactly with the negative to positive sign change of an eigenvalue of (1.15) with $n = 2$.

Generally, it was observed that when symmetric intra-species interactions $g_1 = g_2$ are stronger near the origin than the inter-species interaction g_3, mixing of the

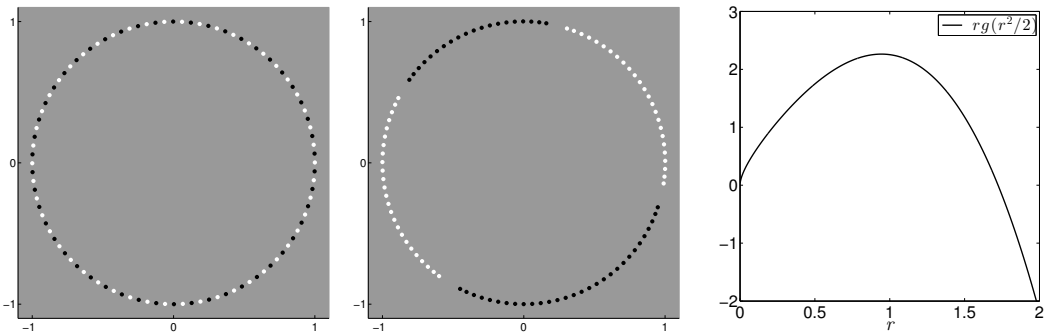


Figure 1.4: Left: An alternating particle ring. Forces $g_1(s) = g_2(s) = 1+2(1-s)+s^{-1/4} - 0.9357796257$, $g_3(s) = 0.5g_1(s)$. Center: Separated particle ring. Forces g_1 and g_2 are the same as on the left, but $g_3 = 1.01g_1$. Right: the true, physical force $rg_1(r^2/2)$.

species tends to occur; when g_3 is stronger, separation tends to occur. See figure 1.9. Self-recognition of species (figure 1.5) may be considered a particular type of separation, and inasmuch as it can be characterized by a mode one instability it can be predicted using theorem 1.6.1.

1.7.2 Instabilities of Low-frequency Modes

Figure 1.5 shows examples of mode five and mode one instabilities. The alternating, symmetric mode five arises from interaction forces defined in terms of

$$\begin{aligned} G3(s) &= 1 + (1-s) + (1-s)^2, \\ G5(s) &= \frac{3}{2}(1-s)^2 + (1-s)^3 - (1-s)^4, \\ G0(s) &= 1 + 2(1-s) + s^{-1/4} - 0.9357796257, \end{aligned} \tag{1.22}$$

as

$$\begin{aligned} g_1(s) &= G5(s) + 1.3 \cdot G0(s), \\ g_2(s) &= G5(s) + 1.3 \cdot G0(s), \\ g_3(s) &= 0.2 * G0(s). \end{aligned}$$

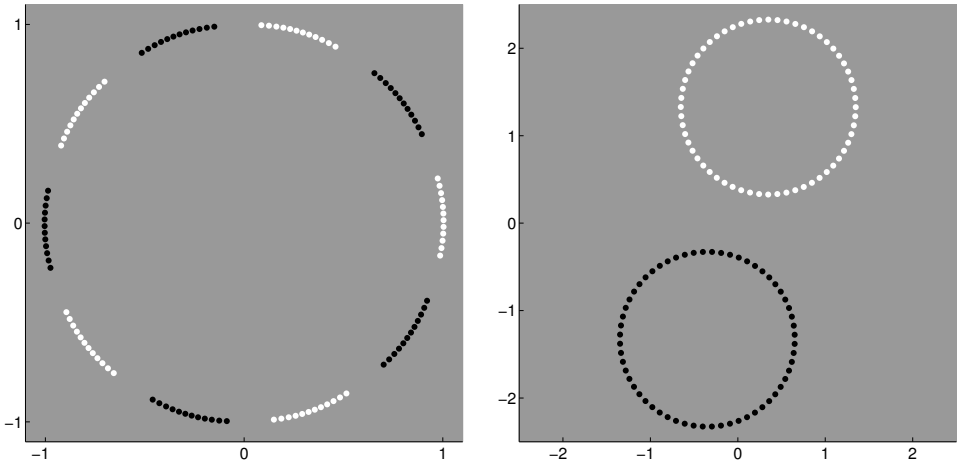


Figure 1.5: Left: symmetric mode five instability. The eigenvector of **E** (1.15) with positive eigenvalue is $[0.02, -0.71, -0.02, 0.71]^T$; the positive-negative $[a, b, -a, -b]$ structure corresponds to the symmetric steady state observed, where the species II density is perturbed with the opposite sign as the species I density. Right: mode one instability.

The mode one instability is due to the interaction forces defined as

$$g_1(s) = G0(s),$$
$$g_2(s) = G0(s),$$
$$g_3(s) = 10^{-4}(C_r e^{-\sqrt{2s}/l_r} - C_a e^{-\sqrt{2s}/l_a}),$$

where $G0$ is from (1.22) and $g3$ Morse with $C_a = 1, l_a = 5, C_r = 4, l_r = 0.5$.

Coupling effects of the rings on each other can be seen in figure 1.6, which shows coupling between type I particles (white) with a mode three instability and type II particles (black) with mode five. The interaction forces are defined in terms of (1.22) as

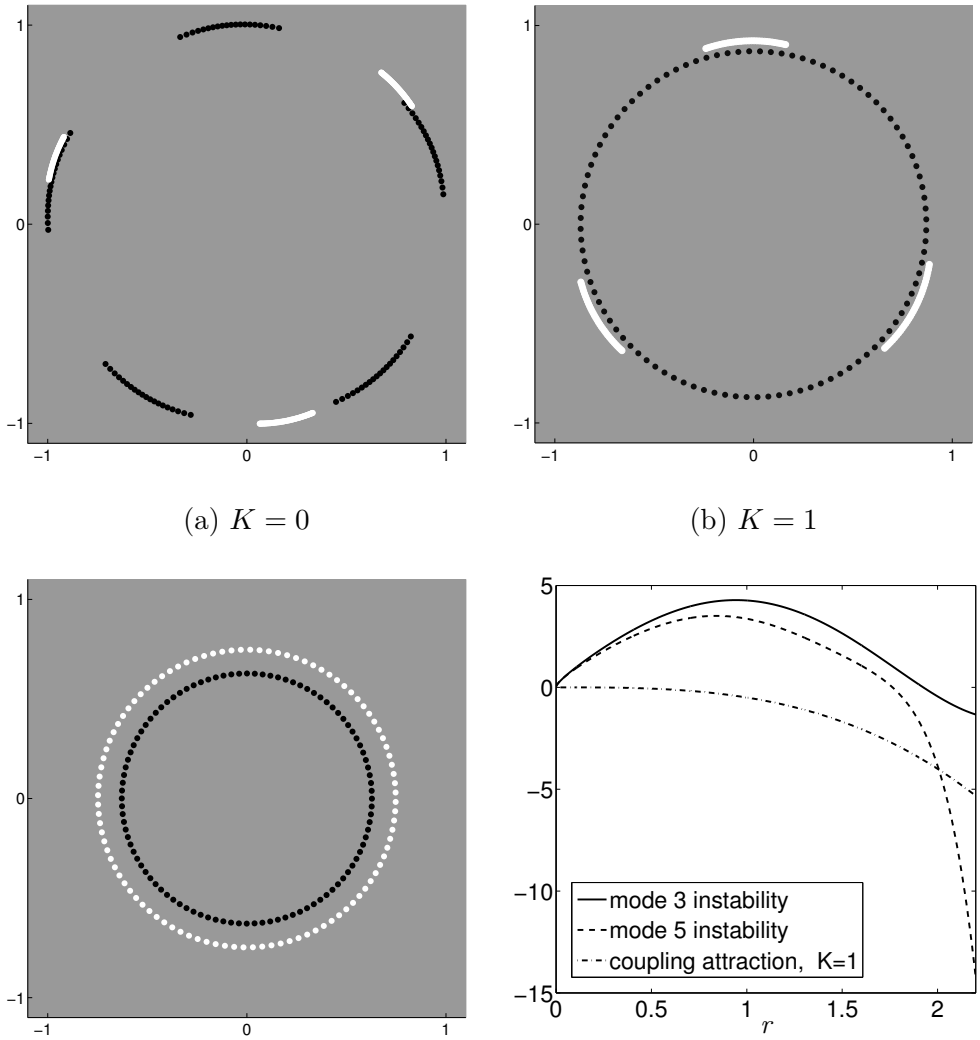
$$g_1(s) = G3(s) + 1.1158 \cdot G0(s),$$
$$g_2(s) = G5(s) + 1.3 \cdot G0(s),$$
$$g_3(s) = -Ks,$$

with $K = 0$, 1, and 4, designed in [86] to exhibit pure mode 3 and 5 instabilities. The sequential disappearance of the instabilities in figure 1.6 corresponds to the eigenvalues of those modes becoming negative (i.e. eigenvalues from (1.15)), confirmed numerically.

Even when the concentric ring solution is not linearly well-posed, the stability or instability of low frequency modes is still borne out in the ground state. Figure 1.7 shows the occurrence of a mode two instability in such a scenario, which is correctly predicted by the linear stability theory even though the linearized equation is ill-posed. When multiple modes become unstable, it is possible that one, several, or all of the unstable modes will appear in the ground state. The question of those which do, and to what degree, is determined by the particular nonlinearity of the problem and is outside the scope of the linear theory. For the single species case, this was pointed out in [86] and the problem of which modes appear is still open. At this time, even less is known about the two-species problem. Weakly nonlinear analysis, considered in [79], may prove useful for this purpose and is one of several considerations for future work.

1.7.3 Linear Ill-posedness

There are several factors which limit the effectiveness of the linear theory when the stable ground state is far from the concentric ring solution. When $R_1 = R_2$, as is the case in many examples with $g_1 = g_2$, the inter-species interaction g_3 must be $o(s^{-1/2})$ as $s \to 0$ for the integrals in 1.15 to exist. This condition is not met in any of the cases of figure 1.1, and so the theory may not be applied. It is possible to replace $g_1(s)$ by $g_1((1+\epsilon)s)$, and if ϵ is sufficiently small then the observed steady state is qualitatively indistinguishable from the unperturbed version but R_1 and R_2 are no longer equal. The theory does apply to this perturbed problem, but another difficulty arises: most or all modes are unstable.



(a) $K = 0$

(b) $K = 1$

(c) $K = 4$

(d) Physical forces

Figure 1.6: Modes three and five stabilize each other as cross-particle attraction increases. Bottom right: true forces corresponding to g_1, g_2, and g_3 with $K = 1$. Changing K scales the coupling force due to g_3.

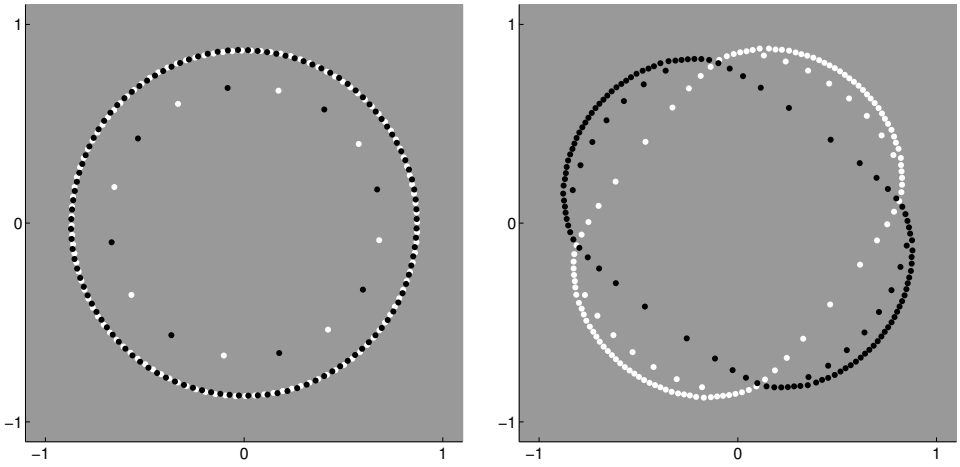


Figure 1.7: Mode two becomes unstable for the power law potential. Left: $g_1(s) = g_2(s) = s^{-0.15}$, $g_3(s) = -s^{0.15}$. Right: $g_1(s) = g_2(s) = s^{-0.15}$, $g_3(s) = -s^{0.2}$.

This type of total instability also occurs in both panels of figure 1.2. While the instability of all low modes is certainly consistent with the observed steady states, it is largely uninformative. In the majority of cases when a mode 1 instability appears, all low frequency modes are unstable as well and the stable steady state is sometimes asymmetric (with the gradient flow coming to rest at a local minimum). Nevertheless, the linear stability theory may still provide some insight. In the right hand panel of figure 1.2, each low mode has one unstable eigenvector except for mode 3, which has two; however, only one and not a linear combination of both appears in the steady state. Why one and not the other or a combination of both appears is impossible to determine by the linear theory, similar to the problem encountered in [86].

That the theory works very well in the cases of figures 1.4, 1.5, 1.6, and 1.7 but not for figures 1.1 and 1.2 is not surprising– when the stable steady state is far from the concentric ring solution, the linear theory is less likely to apply.

Two-particle minimizers also occasionally break symmetry (in the sense that the steady states for the species I and II particles differ by more than a simple

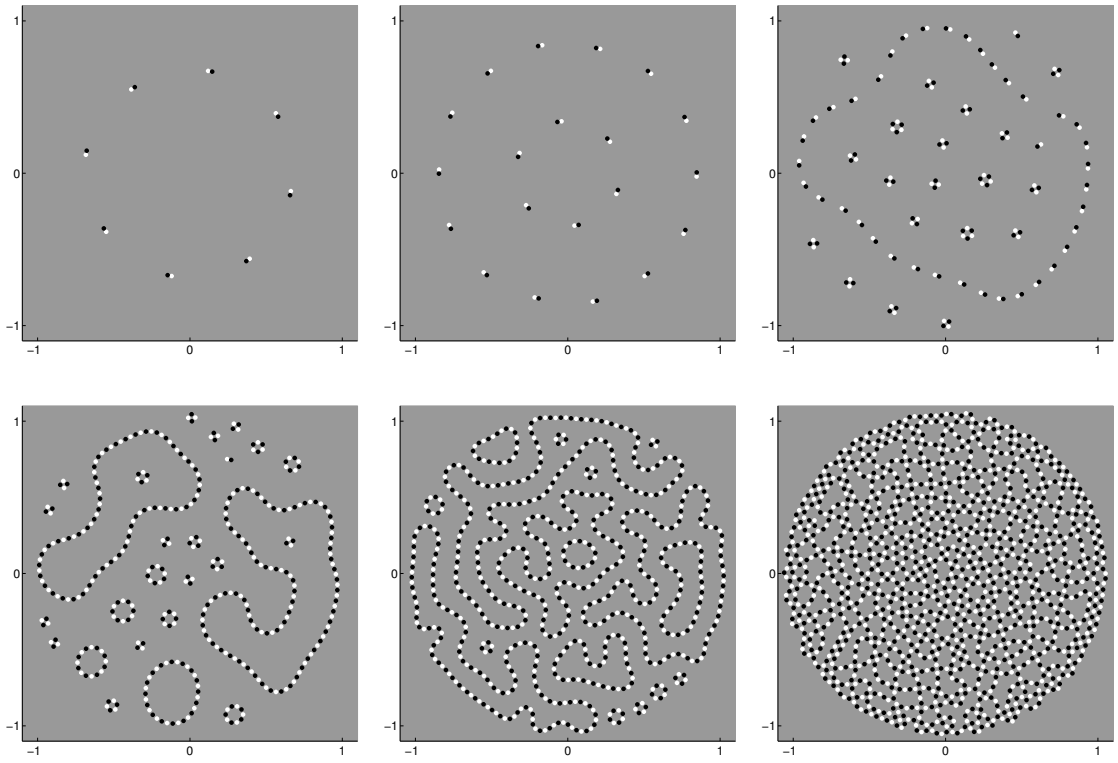
reflection or rotation), pictured in the first panel of figure 1.2. Particle densities from figure 1.1 are asymmetric for certain parameters (see the sixth panel) and may be supported on domains with irregular geometry, including cusps. Some simulations show dependence on initial conditions, which did not manifest for the single species problem. Of course, it is not guaranteed that the gradient flow (1.4), (1.5) reach a global minimizer of the potential (1.3). This minor dependence on initial conditions seems to indicate that the energy landscape for the two-species problem is more complex, and that the gradient flow occasionally comes to rest at local minima.

Simple structures directly relevant to self-assembly, such as alternating particle rings or chains, may also be observed; see figures 1.4 and 1.8. The self-recognition and separation into two rings replicates the phenomena observed in [59], and occurs over a very long time scale ($t \approx 1.4 \times 10^5$) relative to that of the formation of the rings ($t \approx 10^2$).

1.8 Conclusion

Two-species particle aggregation systems have a rich solution structure, including densities with rings, spots, and radial or bilateral N-fold symmetry which often concentrate on or near co-dimension one surfaces. Considering a continuum limit as the number of particles tends to infinity results in a PDE system formulation similar to that of the vortex sheet problem [62, 78]. Linear stability analysis successfully characterizes the steady states which form, verified numerically in section 1.7, and linear well-posedness of the PDE system is considered in section 1.6.

Future work could address the three and higher dimensional versions of the problem, weakly nonlinear analysis of bifurcations from rings to other steady states, the second-order problem, the n-species problem, and the inverse problem



$$V_i(s) = C_{r_i} e^{-\sqrt{2}s/l_{r_i}} - C_{a_i} e^{-\sqrt{2}s/l_{a_i}}$$

$$\begin{array}{lll} C_{a_1} = 1 & C_{a_2} = 1 & C_{a_3} = 1 \\ l_{a_1} = 1 & l_{a_2} = 1 & l_{a_3} = 1.03 \\ C_{r_1} = 2 & C_{r_2} = 2 & C_{r_3} = 1 \\ l_{r_1} = 1 & l_{r_2} = 1 & l_{r_3} = 0.005 \end{array}$$

Figure 1.8: Alternating particle chains arising from the Morse potential, numbers of particles $N_1 = N_2 = N$ with $N = 8, 20, 80, 200, 400, 800$ from top left to bottom right. The first few panels show that the particles seem to form effective dipoles because the inter-species repulsion length scale is so small. When the number of particles increases, the confining nature of the potentials causes them to pack closer together and chains form, as in panel 5. As N increases further, the particles begin to form a two-dimensional lattice structure (panel 6).

of constructing potentials with prescribed instabilities or patterns [86]. In addition, the two-species system allows for the unique possibility of nontrivial H-stable ground states which are outside the scope of the co-dimension one analysis here; c.f. [56].

Figure 1.9 shows a power law example which completely leaves a co-dimension one manifold, and seems to exhibit an effective phase separation or surface tension arising from the nonlocal interactions. As the inter-species repulsion singularity becomes weaker than the intra-species repulsion singularity, black and white particles go from self-segregating to mixing. When the inter-species repulsion is substantially weaker, exhibited in the far right panel of figure 1.9, a regular alternating lattice structure emerges in large portions of the steady state. The formation of lattices in the single species problem is not unfamiliar; see [39] for similar phenomena in the single-species problem. The first two panels of figure 1.9 correspond to local, not global minimizers of the potential 1.3, and a steady state consisting of a straight line interface between the black and white particles has slightly lower energy. The steady states in the right two panels, however, have lower energy than the separated black/white half disks. The analysis carried out in this chapter applies to solutions supported along one-dimensional curves, and as such does not apply to the phenomena in figure 1.9; however, the phase separation effects could be observed on a co-dimension one surface in $\mathbf{R}^3$.

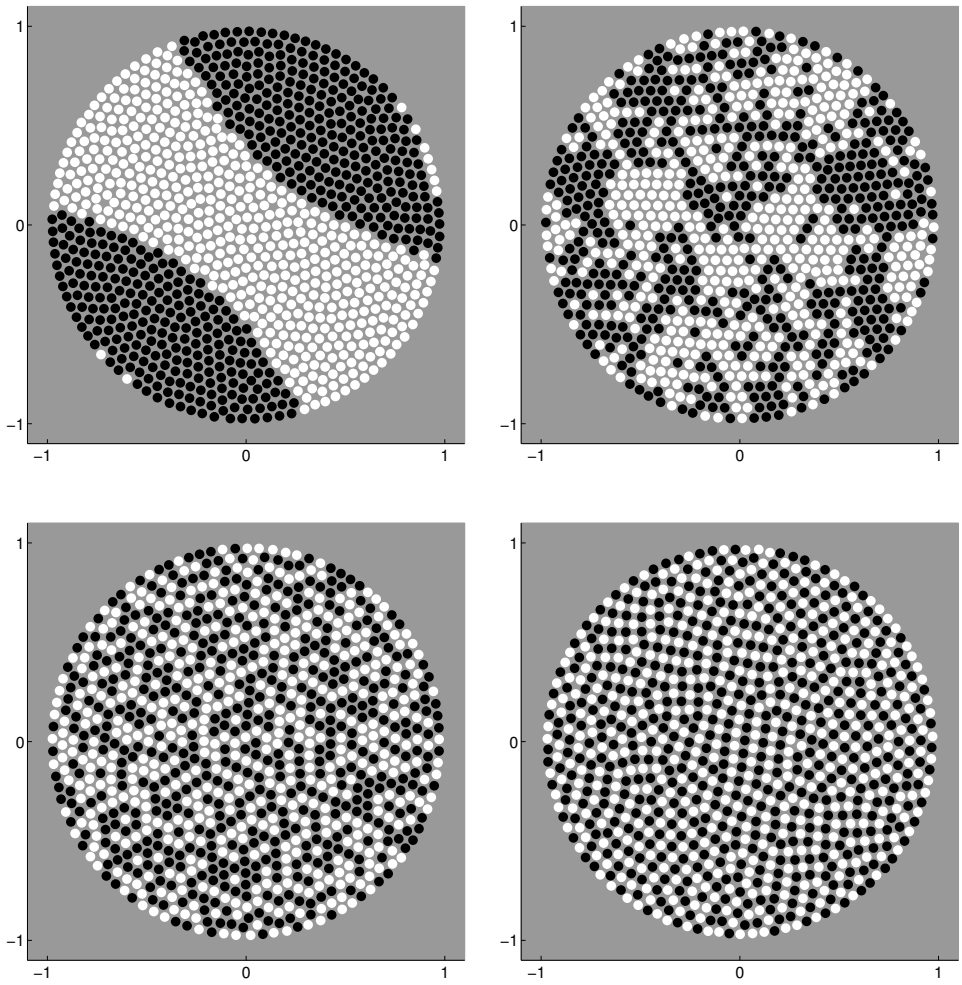


Figure 1.9: Power laws showing phase separation or surface tension as parameters vary. Forces are $g_1(s) = g_2(s) = s^{-1} - 1$, and $g_3(s) = s^{-1+\epsilon} - 1$ with $\epsilon = -0.005, -0.0025, 0.005, 0.025$ from upper left to bottom right.

CHAPTER 2

Sparse Representations for Multiscale PDE

2.1 Preface

This chapter discusses a new area of research (see the references in section 2.2) applying sparse modeling techniques to the numerical solutions of PDE with multiple, separated length scales [61]. A typical example is

$$\frac{\partial u^\epsilon}{\partial t} - \frac{\partial}{\partial x}\left(a(x/\epsilon)\frac{\partial u^\epsilon}{\partial x}\right) = 0 \quad \text{on } [0, 2\pi] \text{ periodic},$$
$$u^\epsilon(x,0) = u_0^\epsilon(x), \quad a(x) \text{ oscillatory}$$

where ϵ is near zero. The highly oscillatory coefficient $a(x/\epsilon)$ introduces ϵ-scale behavior into the solution.

The reason for introducing sparse models is to take advantage of efficient algorithms available for sparse data structures. The results presented here constitute advancements in numerical analysis, but a basic background in sparse modeling and computation with sparse data will help put them in context. This preface provides that background.

2.1.1 Sparse Representations and Sparse Modeling

A vector $x \in \mathbb{R}^n$ is said to be *sparse* if most of its entries are zero. A signal (or image or general data point) $b \in \mathbb{R}^m$ is said to admit a *sparse representation* with respect to a possibly overcomplete basis (or *dictionary*) $\mathbf{D} \in \mathbb{R}^{m\times n}$ if

$$\mathbf{D}x \approx b.$$

This is typically meant in the L^2 sense, i.e. $\|\mathbf{D}x - b\|_2$ is small. A *sparse model* of a signal or set of signals is the prior knowledge that the signals admit a sparse representation with respect to some known dictionary $\mathbf{D}$, or that some transformation of the signals is sparse. Cases in which the signals have a sparse representation are called *synthesis models*; those in which a transformation of the signals is sparse are called *analysis models*.

Sparse models are useful because they provide a powerful form of prior knowledge (this statement is made precise in Bayesian terms at the end of this section) which is crucial for ill-posed inverse problems. Sparse models are particularly powerful because they are widely applicable (many types of data, including images and signals, have sparse representations) and because they can be used to formulate the solutions to ill-posed inverse problems involving this data as minimizers of energy functionals which are computationally tractable.

To illustrate this point, consider the problem of removing noise from a signal which is assumed to have some unknown sparse representation x with respect to a known dictionary $\mathbf{D}$. That is, the true signal is equal to $\mathbf{D}x$ for some sparse x. If the noise is denoted by ϵ, then we observe $b := \mathbf{D}x + \epsilon$ and wish to recover $\mathbf{D}x$.

The sparse model is what makes this possible– without it, the problem is completely underdetermined. We expect that the noise is not too large, so that

$$\|\epsilon\|_2^2 = \|\mathbf{D}x - b\|_2^2 \quad \text{is small,} \tag{2.1}$$

and also that x is sparse:

$$\|x\|_0 \quad \text{is small,} \tag{2.2}$$

where $\|x\|_0$, the L^0 'norm', denotes the number of nonzero entries of x. (Note that it is not a norm.)

We can express our desire to find a representation x satisfying both requirements (2.1) and (2.2) by solving an optimization problem of the form

$$x = \arg\min_x \|\mathbf{D}x - b\|_2^2 \quad \text{s.t. } \|x\|_0 \leq N$$

or

$$x = \arg\min_x \|x\|_0 \quad \text{s.t. } \|\mathbf{D}x - b\| \leq \delta.$$

Alternatively, the search for x can be cast as an unconstrained problem

$$x = \arg\min_x \lambda \|x\|_0 + \frac{1}{2} \|\mathbf{D}x - b\|_2^2 \tag{2.3}$$

where λ which is a parameter. Solving (2.3) uses the L^0 term as a regularizer to promote sparsity of x, and balances this against the reconstruction error $\mathbf{D}x - b$.

The problem with (2.3) is that the L^0 'norm' is totally discontinuous and so the resulting optimization problem is combinatorial in nature and NP-hard [64]. There exist efficient greedy algorithms for solving (2.3) with performance guarantees [84], but these are necessarily approximate in nature. Additionally, it is not entirely clear that the L^0 norm is the best way of measuring sparsity: its discontinuity, at least, is undesirable.

An alternative to approximately solving (2.3) is to consider its convex relaxation

$$x = \arg\min_x \ \lambda \|x\|_1 + \frac{1}{2} \|\mathbf{D}x - b\|_2^2 . \tag{2.4}$$

The advantage of this approach, proposed in [21], is that (2.4) can be solved exactly with efficient convex optimization methods [40]. Another possibility is to replace the L^1 norm by a nonconvex regularizer [19, 20], which often forfeits the guarantee of finding the global minimizer but performs very well in practice. For the purposes of this introduction, we will use the L^1 norm with the understanding that alternatives are available.

This approach can be extended to handle deconvolution, upsampling (or superresolution), and other inverse problems by including the appropriate forward operators into the formulation of (2.4); see the section on TV image reconstruction below. All that is required is that a reasonable optimization problem can be posed with balances a sparsity term ($\|\cdot\|_1$, say) with a data fidelity term (usually $\|\cdot\|_2^2$).

Typical examples of sparse models are given below:

TV/ROF Image Model [75] This is the only analysis model discussed here. The prior knowledge is that images (particularly medical images, such as MRI or X-ray CT scans) are approximately piecewise constant. That is, their gradients are sparse. An image x can thus be approximately recovered from a noisy observation y via

$$x = \arg\min_x \ \lambda \|\nabla x\|_1 + \frac{1}{2}\|x - y\|_2^2 .$$

If the image is also affected by an operator A (which could correspond to blur, downsampling, or more), the optimization problem is

$$x = \arg\min_x \ \lambda \|\nabla x\|_1 + \frac{1}{2}\|Ax - y\|_2^2 .$$

Transform-Domain Models [21] and Sparse Coding [67] These models assume the data b consists of an image or image patch which has a sparse representation x with respect to either a wavelet/chirplet/warplet/shearlet/curvelet/etc dictionary or 'learned' dictionary optimized with respect to a given set of signals. In this case, x represents b in a transform domain defined by the dictionary $\mathbf{D}$. The optimization problem generated for denoising is exactly (2.4).

Morphological Component Models [77] In many cases, it is desirable to separate an observed signal $z = b_1 + b_2$ into its constituent components b_1 and b_2. Examples include the separation of speech from impulse noise and cartoon/texture image decomposition.

Assuming sparse models $b_1 \approx \mathbf{D}_1 x_1$ and $b_2 \approx \mathbf{D}_2 x_2$, the components can be separated by solving the optimization problem

$$x_1, x_2 = \arg\min_{x_1,x_2} \ \lambda_1 \|x_1\|_1 + \lambda_2 \|x_2\|_1 + \frac{1}{2}\|\mathbf{D}_1 x_1 + \mathbf{D}_2 x_2 - z\|_2^2 .$$

Compressed Sensing [17, 18] This celebrated field combines the previously described sparse signal models and efficient convex optimization algorithms with harmonic analysis and random matrix theory to give *provable* guarantees about the recovery of transform-domain sparse signals from undersampled measurements. In this case, the transform is required to be orthonormal (such as a number of wavelet and discrete Fourier transforms).

The setup is as follows: we suppose that $f \in \mathbb{R}^n$ is a signal with a sparse representation in an orthonormal basis $\mathbf{\Psi}$, so that $\mathbf{\Psi}x = f$ with x sparse. We observe $m < n$ linear functional observations of the form

$$b_i = \langle \phi_i, f \rangle = \langle \phi_i, \mathbf{\Psi}x \rangle$$

and wish to recover f. The Nyquist-Shannon sampling theorem guarantees that, in general, this is not possible; we are attempting to find one of infinitely many solutions of an underdetermined linear system $\mathbf{A}f = b$. However, the sparse model $\mathbf{\Psi}x = f$ provides the additional information necessary to reconstruct f.

In other words, defining $\mathbf{A}$ as the matrix with $\phi_1^T \dots \phi_m^T$ as its rows, we wish to find x such that

$$\mathbf{A}\mathbf{\Psi}x = b \quad \text{with } x \text{ sparse.}$$

Along the lines of the other sparse models in this chapter, we propose $f = \mathbf{\Psi}x$ where

$$x = \arg\min_x \ \lambda \|x\|_1 + \frac{1}{2} \|\mathbf{A}\mathbf{\Psi}x - b\| .$$

The point which distinguishes compressed sensing from other sparse modeling contexts is that, under certain conditions on $\mathbf{A}$, solving the above optimization problem is *guaranteed* to succeed. If $\mathbf{A}$ has Gaussian iid entries and x has S nonzero entries, then the number of rows (or measurements) required is just $m = O(S \log n)$ [17].

Parsimonious Statistical Models [81] and Machine Learning L^1 regularization appeared in [81] as a method for enforcing the prior knowledge that most of the coefficients in a linear regression should be zero.

The standard form of a linear regression is to assume that an output y sampled in n instances $y[i]$, $i = 1 \dots n$, depends on inputs $x_1[i] \dots x_p[i]$ in a form which is linear in the parameters $\beta_1 \dots \beta_p$:

$$y[i] = \beta_1 x_1[i] + \ldots + \beta_p x_p[i] + \epsilon[i] \quad \text{for all } i$$

where the $\epsilon[i] \sim N(0, \sigma^2)$ are iid zero-mean Gaussian random variables. In matrix form, the model can be written

$$y \approx \mathbf{X}\beta$$

where the j^{th} column of $\mathbf{X}$ is x_j, $j = 1 \dots p$. Maximum likelihood estimation for this problem leads to [43]

$$\beta = \arg\min_\beta \ \frac{1}{2} \|y - \mathbf{X}\beta\| = (\mathbf{X}^T\mathbf{X})^{-1}(\mathbf{X}^T y).$$

However, if $n < p$ the situation is in general hopeless because the data y could be perfectly explained by infinitely many parameter choices β. Even if $n > p$, there is a danger of overfitting: it may be reasonable to expect that many of the β_j are in fact zero because the corresponding variables x_j have only a negligible effect on y, but the maximum likelihood solution will have $\beta_j \neq 0$ and use these x_j to compensate for some of the error ϵ.

The prior knowledge that many components of β are zero can be incorporated by modifying the optimization problem so that

$$\beta = \arg\min_\beta \ \lambda \|\beta\|_1 + \frac{1}{2} \|y - \mathbf{X}\beta\|,$$

known as LASSO [81]. This use of L^1 regularization is ubiquitous in modern data mining, especially in cases where there are thousands of input variables with negligible effects. The LASSO eliminates them automatically with the tuning of a single parameter λ.

In [91], the authors use the above formulation for face recognition. The x_j are example (cropped and normalized) faces from a database, and y is a face to be recognized. The coefficients β are then either thresholded or used as inputs to a support vector machine or other classifier to determine which individual in the database the face corresponds to.

Outlier Models and Low-Rank + Sparse Decompositions [16] In many contexts, such as collaborative filtering and topic modeling, it is assumed that a matrix $\mathbf{Y} \in \mathbb{R}^{n \times m}$ of observations can be approximately factored

$$\mathbf{Y} \approx \mathbf{UV}$$

where $\mathbf{U} \in \mathbb{R}^{n \times k}$ and $\mathbf{V} \in \mathbb{R}^{k \times m}$ with k much smaller than either m or n.

However, it may be the case that a relatively small number of entries of Y do not fit the low-rank model. Along the lines of the morphological component model discussed previously, we can formulate

$$\mathbf{Y} \approx \mathbf{UV} + \mathbf{S}$$

where the majority of entries of $\mathbf{S}$ are zero. This leads to the optimization problem

$$\mathbf{U}, \mathbf{V}, \mathbf{S} = \arg\min_{\mathbf{U},\mathbf{V},\mathbf{S}} \ \lambda \|\mathbf{S}\|_1 + \frac{1}{2} \|\mathbf{UV} + \mathbf{S} - \mathbf{Y}\|_2^2 .$$

Other regularizers, such as the nuclear norm [72], can be used to limit the rank of a matrix directly rather than resorting to the factorization $\mathbf{Y} \approx \mathbf{UV}$. The advantage is that using the nuclear norm leads to a convex optimization problem rather than the nonconvex one above. The low-rank + sparse decomposition finds applications in video foreground-background separation and other areas [16].

Aside: Sparsity as a Bayesian Prior

In the above, the term "prior" has been used loosely to describe additional knowledge used to solve ill-posed inverse problems, separate a signal into components,

and seek meaningful solutions to large linear regression problems. Here, we show that this usage coincides with the way the term is used in Bayesian statistics.

To illustrate this point, we consider the problem of recovering a sparse vector $x \in \mathbb{R}^n$ from noisy measurements $b = \mathbf{A}x + \epsilon$, where $\mathbf{A} \in \mathbb{R}^{n \times m}$ and ϵ is a vector of mean-zero Gaussian noise with variance σ^2. We specify a Laplace prior

$$p(x) = C_1 \exp\{-\|x\|_1 / \lambda\}$$

which is the most common sparsity prior for x, but there are other choices. C_1 is a normalizing constant.

The likelihood is Gaussian:

$$p(\mathbf{D}x - b \,|\, x) = C_2 \exp\left\{-\frac{1}{2\sigma^2} \|\mathbf{D}x - b\|_2^2\right\},$$

and so the posterior is given by Bayes' rule as

$$p(x|b) \propto p(b|x)p(x) \propto C_1 C_2 \exp\left\{-\|x\|_1 / \lambda - \frac{1}{2\sigma^2} \|\mathbf{D}x - b\|_2^2\right\}.$$

The posterior attains its maximum at

$$x^* = \arg\min_x \ \frac{\sigma^2}{\lambda} \|x\|_1 + \frac{1}{2} \|\mathbf{D}x - b\|_2^2$$

which is exactly the optimization problem proposed above. In Bayesian terms, using this value of x at the mode of the posterior is called maximum a-posteriori (MAP) estimation. A true Bayesian treatment would avoid using a point estimate for x at all, preferring the whole posterior distribution instead. When a point estimate is necessary, the usual Bayesian choice is the mean of the posterior instead of its maximum. Computing the posterior mean is usually done with Markov Chain Monte Carlo (MCMC) methods, or variational inference.

2.1.2 Computation with Sparse Data

The application to numerical PDE we are considering does not involve any ill-posed problems for which sparsity provides the necessary additional information

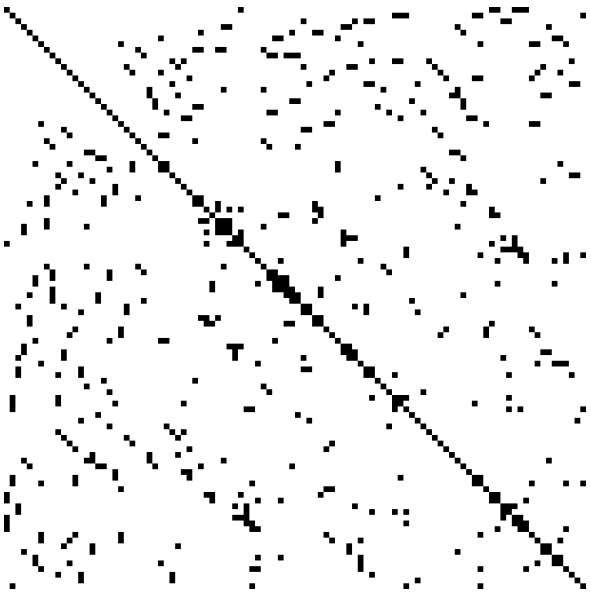


Figure 2.1: Sparsity pattern of a typical finite element matrix. Only the black pixels correspond to nonzero entries. Figure courtesy of Wikipedia [88].

to recover a solution. The PDEs we consider are well-posed, and have unique solutions.

Instead, the primary role of sparsity in this chapter is to allow for efficient computation. This is already common practice in numerical analysis and scientific computing, and it is one of the primary advantages of iterative methods such as Conjugate Gradient, MINRES, and GMRES [41].

The basic idea is that, when most of the entries in a matrix or vector are zero, memory and time can be saved by storing only the location and value of the nonzero entries. For example, matrices arising in finite element methods have the property that the number of nonzero entries in each row (and column) is bounded by the maximum number of elements adjacent to any given element, which stays fixed as the mesh is refined. Figure 2.1 shows such a matrix.

If a sparse matrix $\mathbf{A}$ is size $N \times N$, naive multiplication against an ordinary dense vector requires $O(N^3)$ floating point operations. However, because $\mathbf{A}$ is sparse, there exist, say, only k nonzeros per row. By only multiplying the nonzero

entries of A times a vector, the number of floating point operations can be reduced to $O(kN)$. Since k stays fixed as N increases, this is a huge complexity gain.

2.1.2.1 Efficient Convolution of Sparse Vectors

This chapter focuses on solving PDEs of the form

$$\frac{\partial u^\epsilon}{\partial t} - \frac{\partial}{\partial x}\left(f(x/\epsilon)\frac{\partial u^\epsilon}{\partial x}\right) = 0 \quad \text{on } [0, 2\pi] \text{ periodic,}$$
$$u^\epsilon(x,0) = u_0^\epsilon(x), \quad f(x/\epsilon) \text{ oscillatory}$$

in the Fourier domain. Explicit methods and implicit methods based on Krylov subspace solvers all require repeatedly applying the elliptic operator

$$\frac{\partial}{\partial x}\left(f(x/\epsilon)\frac{\partial u^\epsilon}{\partial x}\right),$$

which comes down to convolutions of the form

$$x * y,$$

where x and y are sparse, length-N vectors representing $\hat{f}$ and $\hat{u}$ discretized.

If x has n_x nonzero entries and y has n_y, then

$$(x * y)[i] = \sum_{j=0}^{N-1} x[i - j \bmod N] \cdot y[j]$$

can be thought of as the summation of n_y scaled shifts of x, so that computing the convolution is tantamount to adding together n_y sparse vectors (it can also be thought of as a matrix-vector multiplication). We assume that x and y are stored as lists of (index, value) pairs, the most commonly used format for sparse vectors. Specifically, we assume there is a queue data structure q_x which represents x. That is, q_x has n_x entries $\{q_x[k]\}_{k=0}^{n_x-1}$, each of which is an $(index, value)$ pair $(q_x[k].first, q_x[k].second)$ such that $x[q_x[k].first] = q_x[k].second$. We assume that the queue allows iteration over its elements, and $O(1)$ insertion at the back, and $O(1)$ deletion from the front (these are all standard assumptions for queues). For

example, such a queue could be implemented easily with a linked list. We assume q_y is the analogous queue representing y.

For example, if x is a vector of length 10 with three nonzero entries $x[1] = -1$, $x[5] = 1$, and $x[7] = 6$, the queue q_x would consist of the pairs $(1, -1)$, $(5, 1)$, $(7, 6)$.

We now discuss several possible algorithms for computing the convolution of two sparse vectors, which is the main bottleneck in our numerical procedure.

Merge Approach If we assume that the queues q_x and q_y representing x and y are sorted by increasing index (as in the above example), the problem can be viewed as one of merging sorted lists into a longer sorted list. When there is a 'tie', i.e. two of the vectors to be added have a nonzero in the same index, the corresponding values are simply added in the result.

Naively merging the sorted lists by searching for the minimum index among the n_y shifts of x would result in an algorithmic complexity of $O(n_y^2 n_x)$. At each iteration, the front entry of each of the n_y shifts would be checked to determine the minimum, a $O(n_y)$ cost. Each iteration adds only one additional coefficient to the result, so $n_x n_y$ total iterations are required, leading to the overall $O(n_y^2 n_x)$ complexity. This has the advantage of being independent of the total grid size N, but as it turns out this algorithm is far from optimal.

Keeping track of the minimum index during the merge can be accomplished more efficiently with a heap data structure [89], which is ideally suited for this purpose. Details of the algorithm revolve around data structures rather than mathematics, and a more efficient algorithm will be presented next, so pseudocode for this approach is omitted. The complexity is $O(n_x n_y \log(n_y))$, a substantial improvement over the naive algorithm.

Index-tracking Approach Ideally, it would be possible to compute the convolution (2.1.2.1) with $O(n_x n_y)$ complexity. This is the best possible, because the sum in (2.1.2.1) consists of $n_x n_y$ terms which each require a floating point multiplication.

It turns out that this lower bound is achievable, and the key to reaching it is to notice that the extra $\log(n_y)$ factor in the complexity of the merge approach is due to the computational cost of keeping track of the minimum index. However, this can be avoided in two ways. The first is to use the fact that the vectors to be added are not arbitrary, but rather shifts of a single vector x. Therefore, the vectors which will contribute to a nonzero entry in a particular index of the result can be known ahead of time rather than computed with the heap.

The second approach, which is more general, is to use an auxiliary vector to store the result with an additional array to keep track of which entries are nonzero. This way, the indices of the nonzero entries do not need to stay sorted. Our pseudocode for the algorithm assumes the following notation:

- q_x is a queue representing x in (index, value) pairs as described above.
- q_y is an analogous queue representing y.
- z is an empty queue of the same form, which will store the result of the convolution.
- ind_list is an empty queue, which will hold integers in the range $0 : N-1$. Let $ind\_list.length$ denote its length.
- t is an array with entries $\{t[i]\}_{i=0}^{N-1}$. Recall that arrays offer $O(1)$ entry lookup, which will be crucial later.

Pseudocode for the sparse convolution - index-tracking algorithm is given below, and accomplishes $MAX\_ITERATIONS$ sparse convolutions with shrink.

Sparse Convolution - Index-tracking Algorithm

initialize $t[i] = 0$ for all $i = 0 : N - 1$ ▷ $O(N)$ cost

for $counter = 0 : MAX\_ITERATIONS - 1$ **do**

  for $j = 0 : n_y - 1$ **do**

    for $k = 0 : n_x - 1$ **do**

      if $(t[q_x[k].first + q_y[j].first \bmod N] = 0)$ **then**

        append $q_x[k].first + q_y[j].first \bmod N$ to ind_list ▷ $O(1)$ cost

      end if

      t$[q_x[k].first + q_y[j].first \bmod N] \mathrel{+}= q_x[k].second \times q_y[j].second$ ▷ $O(1)$ cost

    end for

  end for

  for $i = 0 : ind\_list.length - 1$ **do**

    if $(|t[ind\_list[i]]| \geq \lambda)$ **then**

      append $(ind\_list[i], \text{shrink}(t[ind\_list[i]]))$ to z ▷ $O(1)$ cost

    end if

    set $t[ind\_list[i]] = 0$ ▷ $O(1)$ cost

  end for

end for

return z

The nested loops over j and k lead to a total cost of $O(n_x n_y)$, the best one could hope for. At each step of this loop, at most one entry is added to ind_list, so the loop over i costs at most $O(n_x n_y)$ as well. Thus the whole cost is

$$O(n_x n_y MAX\_ITERATIONS) + O(N)$$

where the only $O(N)$ cost is the initialization in the first line, which is cheap.

This algorithm strengthens the results of the chapter by demonstrating that efficient computation, independent of the grid size N after the first iteration, is possible. There are other approaches to multiscale PDE with scale separation (e.g. [28]) which leverage sparse representations of the solution for efficiency, but so far none with algorithmic complexity totally independent of the grid size (except for the initial and final Fast Fourier Transforms).

2.2 Introduction

Partial differential equations with multiple length scales are fundamental to modeling various physical problems including composite materials, wave propagation in inhomogeneous media, crystalline solids, and flows with high Reynolds number (fluid mechanics). Typically, these problems involve a wide range of scales, with each scale corresponding to a level of physical processes. However, in some cases, the problem is scale separable, in the sense that the mathematical representation of the dynamics involve one fine scale and one course scale. Even in this case, accurate numerical methods for solving these PDE can be computationally expensive since resolving both the coarse and fine scales simultaneously requires a spatial resolution dominated by the fine scale.

Over the past decades, various approaches have been taken to overcome this difficulty. In some cases, it is possible to derive an asymptotic approximation for the effect of small scales on the solution [70]. When this is not possible, many other techniques have been proposed. A multiscale finite elements method can

be used to solve linear elliptic homogenization equations (see [47]), and has found many applications to other multiscale problems. The equation-free methods use accurate small scale and short time solvers to capture fine scale behavior and use them to govern the related course scale behavior [49]. The heterogeneous multiscale method [34] is a general numerical approache which also uses the scale separation of the problem to generate solvers on the micro and macroscopic levels. In [66], a projection based approach is used to construct an adaptive multiscale algorithm for elliptic homogenization equations. And more recently, a sparse transform method [28] exploits the scale separability of linear homogenization problems to construct a fast direct solver. The body of literature on multiscale models is large, and we only mention some of the popular methods. For more detail on general numerical methods for multiscale problems see [35, 34] and the citations therein.

In this work, we will focus our attention on linear partial differential equations with multiscale behavior either in the medium or in a source term. Following the work of [76], which used an L^1 optimization method to compress the Fourier coefficients of the solution, we build efficient solvers for periodic multiscale problems. In particular, we will use the sparse Fourier structure of solutions to construct numerical methods which solve the problem directly, without the separating the micro and macro scales explicitly.

L^1 optimization and its related models are at the center of many problems in the fields of imaging science and data analysis, see for example [17, 16, 31, 15]. Due to the connection with sparse models for compressive sensing, recent works have introduced L^1 techniques for numerical partial differential equations. For example, in [76] L^1 regularized least squares was used to sparsely approximate the Fourier coefficients in multiscale dynamic PDE (and in this work we expand that approach). In [65, 68, 69], eigenfunctions with compact support were constructed to efficiently solve problems in quantum mechanics. Also, in [46] an L^1 nonlinear

least squares model was used to sparsely recover coefficients of a second order ODE which are related to constructing intrinsic mode functions. In [11], low-rank libraries are used to sparsely approximate solutions to dynamical systems and thereby identify bifurcation regimes. Some theoretical results are provided in [14] for PDE with L^1-terms, related to some of these models. For more detailed analytic results, see [8, 9, 10] which laid the theoretical groundwork for these equations.

In this chapter, we continue the work of [76] to leverage the sparsity of solutions in order to design an efficient numerical scheme. However, we also impose sparsity of the update operator to improve the complexity while retaining a similar level of accuracy. We show some theoretical results for the sparse spectral scheme and sparse operator-sparse solution spectral scheme. In particular, we provide error bounds between the solution and the sparse approximation as well as complexity bounds on the algorithm. Also, we continue to make connections between L^1 based methods and multiscale problems through a denoising interpretation of the homogenization expansion of the solution.

The outline of this work is as follows. In Section 2.3, we recall the explicit scheme from [76] and in 2.4 propose an implicit version as well as a sparse operator approximation. Theoretical results are provided in Section 2.5. A discussion on well-posedness is given in Section 2.6 and a denoising interpretation of the method is given in Section 2.7. In Section 2.8, some algorithmic analysis is provided. The algorithm is tested on numerical examples in Section 2.9, with concluding remarks given in Section 2.10.

2.2.1 Notation

- a – (or A for anisotropic problems) the medium inhomogeneity. $\hat{a}'$ is the sparse approximation of $\hat{a}$.

- μ – the shrink size variable. μ' is the corresponding variable for sparse operator approximation.
- k – the Fourier space variable, with positive and negative frequencies.
- Q – either a general numerical scheme or the matrix corresponding to a one-step linear numerical scheme.
- L – an elliptic operator. $\hat{L}$ is the operator when applied in the Fourier domain and $\hat{L}_h$ is its discretization. $\hat{L}'_h$ is the sparse approximation.

2.3 Preliminary

We will consider linear multiscale problems where the solutions are sparse in the Fourier domain [28, 76]. For example, consider the parabolic problem:

$$\begin{aligned} \frac{\partial u^\epsilon}{\partial t} - \frac{\partial}{\partial x}\left(a(x/\epsilon)\frac{\partial u^\epsilon}{\partial x}\right) = 0 \quad &\text{on } [0, 2\pi] \text{ periodic} \\ u^\epsilon(x,0) = u_0^\epsilon(x), \quad &a(x/\epsilon) \text{ oscillatory.} \end{aligned} \tag{2.5}$$

Figure 2.2 shows the solution in physical and Fourier space. This phenomenon is common in multiscale PDE: distinct length scales manifest strikingly as sparsity in the frequency domain.

To compute solutions which are truly sparse in the frequency domain (and not just approximately sparse with many noisy small magnitude coefficients), it was proposed in [76] to solve an ℓ^1-regularized least squares problem to obtain a sparse approximation of $\hat{u}$ (the Fourier transform of u). We summarize the method here.

Given numerical iterates $\hat{u}^n, \ldots, \hat{u}^{n-q}$ and a numerical update scheme of the form

$$\hat{u}^{n+1} = Q(\hat{u}^n, \ldots, \hat{u}^{n-q}),$$

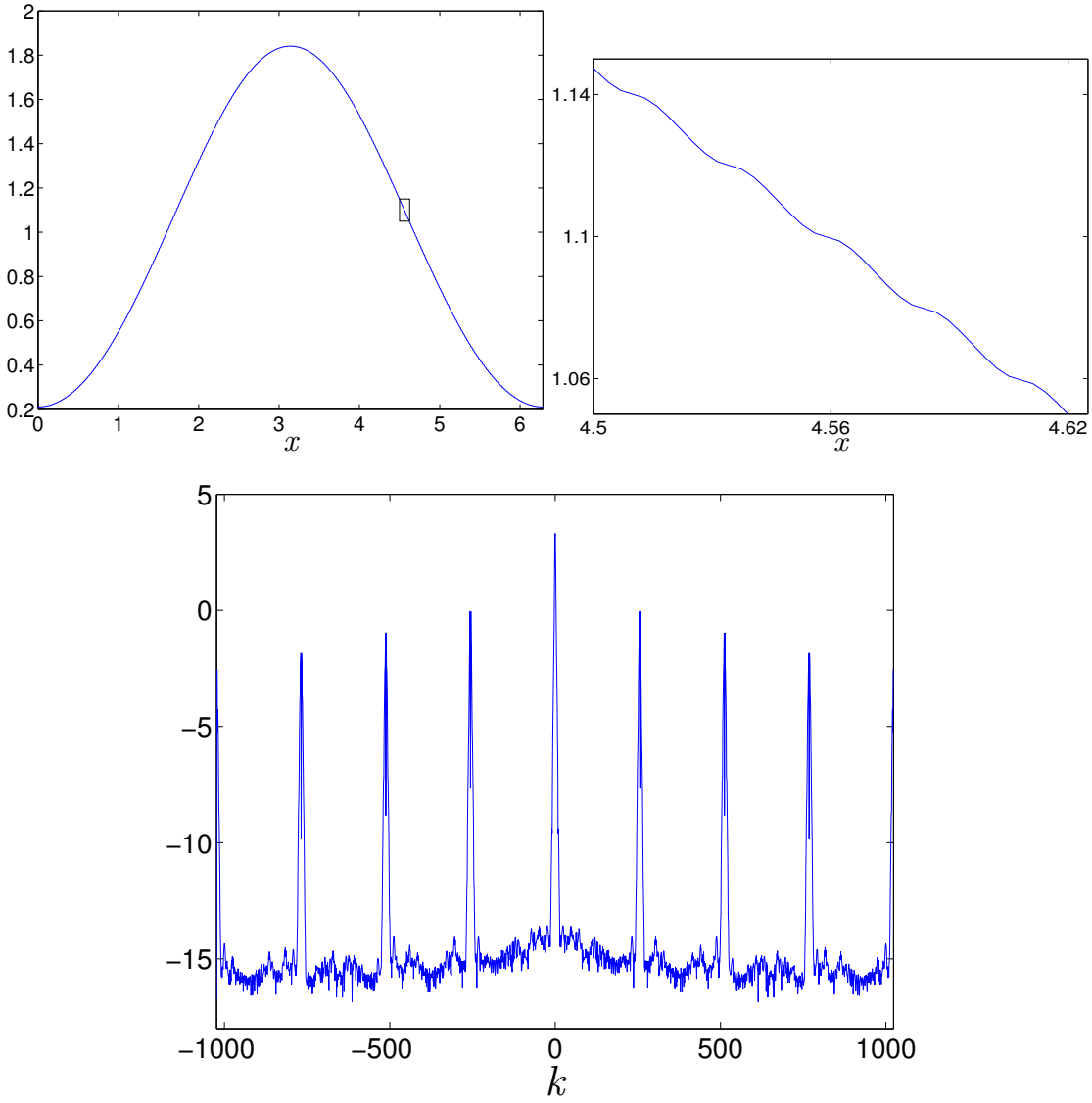


Figure 2.2: **Left:** Solution of (2.5) with Fourier-sparse initial data in physical space. The small rectangle shows the axis limits of the zoomed in plot to the right. **Right:** Zoomed in, showing fine scale oscillations. **Bottom:** solutions in Fourier space (the y-axis for all Fourier space plots is on a $\log_{10}$ scale). Of the $N = 2048$ Fourier coefficients, only 153 have magnitude larger than 10^{-10}.

the scheme is modified by defining the auxiliary variable $\hat{v} = Q(\hat{u}^n, \ldots, \hat{u}^{n-q})$ and solving

$$\hat{u}^{n+1} = \arg\min w \;\; \mu \|w\|_1 + \frac{1}{2} \|w - \hat{v}\|_2^2 , \tag{2.6}$$

where the ℓ^1 norm for complex arguments w is $\|w\|_1 = \sum_i |w_i|$, where $|\cdot|$ denotes magnitude. Note that the ℓ^1 norm is taken in the Fourier domain and not physical space.

For a one-step linear updating scheme, equation (2.6) can be written as

$$\hat{u}^{n+1} = \arg\min w \;\; \mu \|w\|_1 + \frac{1}{2} \|w - Q\hat{u}^n\|_2^2$$

where Q is the matrix which advances the discretized solution forward in time. L^1 regularized least squares is amenable to a number of efficient solution methods, e.g. [40]. The problem can also be generalized to any basis or overcomplete dictionary, but we restrict our attention to Fourier modes. In fact, due to the orthogonality of the Fourier modes, equation (2.6) decouples and the minimizer can be given exactly:

$$\hat{u}^{n+1} = \text{shrink}(\hat{v}, \mu) := \max(|\hat{v}| - \mu, 0) \frac{\hat{v}}{|\hat{v}|}.$$

For a concrete example, the forward Euler method applied to $\frac{\partial u^\epsilon}{\partial t} = \frac{\partial}{\partial x}\left[a(x/\epsilon)\frac{\partial u^\epsilon}{\partial x}\right]$ has the form:

$$\hat{u}^{n+1} = \hat{u}^n + dt\, i\, k\, \hat{a} * (i\, k\, \hat{u}^n)$$

where k is the wave number and $*$ represents convolution. This becomes:

$$\hat{u}^{n+1} = \text{shrink}\,(\,\hat{u}^n + dt\, i\, k\, \hat{a} * (i\, k\, \hat{u}^n), \mu\,)$$

in the sparse spectral form. By exploiting sparsity in the frequency domain, the proposed method can rely on sparse data structures to allow for high resolution with faster numerical simulations.

2.4 Proposed Methods

In this section, we will discuss two new extensions of the sparse spectral method, namely an implicit version and a sparse operator/sparse solution version. Each come with their own advantages, which we will analyze in subsequent sections.

2.4.1 Implicit Variation

For many classes of problems at high spatial resolution, explicit schemes are impractical due to the severe time step restriction required for stability. We can construct an implicit scheme for the problems we are considering, which avoids these restrictions at the expense solving a more complex L^1 problem at each time step.

Consider the general linear parabolic equation $u_t + Lu = f$ with schemes of the form

$$Q\hat{u}^{n+1} = \hat{u}^n + dt\hat{f}_h.$$

The simplest implicit version is backward Euler:

$$(I + dt\hat{L}_h)\hat{u}^{n+1} = \hat{u}^n + dt\hat{f}_h,$$

where $\hat{L}$ denotes the representation of L in the Fourier basis, $\hat{L}_h$ denotes the discretized version of this operator with respect to a grid size $h > 0$, and $\hat{f}_h$ denotes the Fourier transform of f sampled at the corresponding grid points.

For a scheme of this form, the analogue of equation (2.6) is

$$\hat{u}^{n+1} = \arg\min w \;\; \mu \|w\|_1 + \frac{1}{2}\left\|Qw - (\hat{u}^n + dt\hat{f}_h)\right\|_2^2 \tag{2.7}$$

which does not have a simple explicit representation. In addition, the optimality condition for Equation (2.7) requires inverting the matrix Q^TQ which will often be badly conditioned.

When L is a uniformly elliptic operator, the eigenvalues of $Q = I + dt\hat{L}_h$ are positive and so we can instead consider the sparse scheme defined by

$$\hat{u}^{n+1} = \arg\min w \;\; \mu \|w\|_1 + \frac{1}{2} w^T Q w - w^T (\hat{u}^n + dt \hat{f}_h). \tag{2.8}$$

Similarly for time-independent problems, *i.e.* $Lu = f$, the corresponding energy is

$$\hat{u} = \arg\min w \;\; \mu \|w\|_1 + \frac{1}{2} w^T \hat{L}_h w - w^T \hat{f}_h.$$

Note that when $\mu = 0$, this is the standard variational principle for elliptic operators. We will see that solving the implicit schemes with the L^1 term directly is often too slow to be practical. The reason is that directly applying this variational principle to find the solution does not use the fact that the solution is sparse in order to speed up computations. However, in Section 2.8.1 we will show that it is possible to construct an efficient algorithm for approximately solving the resulting optimality condition arising from equation (2.8).

2.4.2 Sparse Operator Approximation

For uniformly elliptic linear operators, for example of the form

$$Lu = -\mathrm{div}(a(x, x/\epsilon)\nabla u),$$

the standard spectral discretization

$$\hat{L}_h \hat{u} = k\, \hat{a} * (k\, \hat{u})$$

requires a convolution at each iteration, which can be costly even when $\hat{u}$ is sparse. However, because the diffusion coefficient a is scale separated, we can define a sparse approximation of $\hat{L}_h$ by

$$\hat{L}'_h \hat{u} = k\, \hat{a}' * (k\, \hat{u})$$

where $\hat{a}'$ is a sparse approximation of $\hat{a}$. We can choose $\hat{a}'$ to solve

$$\hat{a}' = \arg\min w \;\; \mu' \|w\|_1 + \frac{1}{2} \|w - \hat{a}\|_2^2$$

which again results in a closed form solution given by the soft thresholding $\hat{a}' = \text{shrink}(\hat{a}, \mu')$. An alternative is

$$\hat{a}' = \arg\min w \ \mu' \|w\|_0 + \frac{1}{2}\|w - \hat{a}\|_2^2$$

where the L^0 'norm' $\|\cdot\|_0$ counts the number of nonzero entries. In this case, the solution is given by hard thresholding $\hat{a}$— setting all coefficients smaller in magnitude than some threshold equal to zero.

Soft thresholding is contractive and benefits from many desirable smoothing properties which make it preferable for the sparse approximation of the solution, which will be discussed below in Section 2.7. For a sparse approximation of the operator, the benefits of a particular choice of thresholding are less clear and therefore we consider both.

2.5 Theoretical Remarks

The compressive spectral method, or sparse scheme, inherits many appealing properties of the underlying numerical method it approximates. In general, it is at least as stable as the original scheme and retains the order of accuracy.

2.5.1 Contraction and Linear Convergence

The following two theorems show that the explicit and implicit numerical schemes are contractive. This result is similar to those found in [9].

Theorem 2.5.1. *For the explicit scheme generating time steps by*

$$< \hat{u}^{n+1} = shrink((I - dt\hat{L}_h)\hat{u}^n + dt\hat{f}, \mu),$$

if $||I - dt\hat{L}_h||_{op} \leq 1$ *then the iterations are contractive: i.e.* $||u^{n+1} - u^n||_2 \leq ||u^n - u^{n-1}||_2$.

Here $\|\cdot\|_{op}$ denotes the ℓ^2 operator norm, or largest singular value.

Theorem 2.5.2. *For the implicit scheme, if $\hat{L}_h$ is positive semidefinite, then the iterations are contractive, $||u^{n+1} - u^n||_2 \leq ||u^n - u^{n-1}||_2$, for all $dt > 0$.*

The proofs of these two theorems are similar, and reside in the appendix.

The method is also convergent. In particular, for the correct scaling of μ, we have the following theorem.

Theorem 2.5.3 (Linear Convergence, Explicit Scheme)**.** *Let S denote a linear spectral numerical update scheme, generating time steps as*

$$\hat{u}^{n+1} = Q(\hat{u}^n, \ldots, \hat{u}^{n-k}),$$

and let S_μ denote the spectrally sparse scheme, which generates time steps as

$$\hat{u}_\mu^{n+1} = shrink(Q(\hat{u}_\mu^n, \ldots, \hat{u}_\mu^{n-k}), \mu).$$

Then if S is consistent and stable (and hence converges), and if $\mu = O(dt^{1+\delta})$ for some $\delta > 0$, then the compressive scheme S_μ converges. If $\mu = O(dt^p)$ with p at least the order of the local truncation error of S, then the order of convergence of S is not impacted.

For the implicit scheme, the analogous theorem is the following.

Theorem 2.5.4 (Linear Convergence, Implicit Scheme)**.** *Let S denote an implicit linear spectral numerical update scheme for the PDE $u_t + Lu = f$ on a domain Ω discretized with N grid points, generating time steps as*

$$(I + dt\hat{L}_h)\hat{u}^{n+1} = \hat{u}^n + dt\hat{f}_h,$$

and let S_μ denote the spectrally sparse scheme, which generates time steps according to (2.8):

$$\hat{u}^{n+1} = arg\,min\, w \;\; \mu \|w\|_1 + \frac{1}{2} w^T (I + dt\hat{L}_h)w - w^T(\hat{u}_\mu^n + dt\hat{f}_h).$$

Then if S is consistent and stable (and hence converges), and if $\mu = O(dt^{1+\delta})$ for some $\delta > 0$, then the spectrally sparse scheme S_μ converges. If $\mu = O(dt^p)$ with p at least the order of the local truncation error of S, then the order of convergence of S in L^2 is not altered.

The proofs can be found in the Appendix.

2.5.2 Sparse Operator Approximation: Implicit Solver

We now consider the error incurred by the sparse operator approximation proposed in Section 2.4.2. The continuum case is discussed in detail, and the proof for the case of discretized operators is completely analogous.

The usual discretization in Fourier space of a general, anisotropic, divergence form elliptic operator

$$Lu = -\text{div}(A(x)\nabla u) + b(x) \cdot \nabla u + c(x)u$$

results in a matrix (corresponding to convolution) which is dense. However, it is still approximately sparse when the coefficients A and b are. Approximating A and b by A' and b' which are truly sparse in Fourier space yields an operator which is far more efficient to store and to work with, but incurs some error. Theorem 2.5.5 quantifies this error.

Theorem 2.5.5. *Let u_1 and u_2 be solutions to*

$$-\text{div}(A_1\nabla u_1) + b_1 \cdot \nabla u_1 + c_1 u_1 = f \tag{2.9a}$$

$$-\text{div}(A_2\nabla u_2) + b_2 \cdot \nabla u_2 + c_2 u_2 = f. \tag{2.9b}$$

on a domain $\Omega \subset \mathbb{R}^d$ with periodic boundary conditions and the constraint

$$\int_\Omega u_1 = \int_\Omega u_2 = 0.$$

Require also that

$$w^T A_i w \geq \lambda \|w\|^2,$$
$$c_i - \frac{1}{2}\mathrm{div}(b_i) \geq 0$$

for $i = 1, 2$. *Then*

$$\|u_1 - u_2\|_{H^1} \leq C\lambda^{-2}\Bigg(d \max_{i,j} \left\|(\hat{A}_1)_{ij} - (\hat{A}_2)_{ij}\right\|_1 + C \max_i \left\|(\hat{b}_1)_i - (\hat{b}_2)_i\right\|_1 + C^2 \|\hat{c}_1 - \hat{c}_2\|_1 \Bigg) \|f\|_2$$

where $C = C(\Omega)$ *is the constant from Poincare's inequality [38], and the Fourier series of the matrices and vectors* A_i *and* b_i *are taken entry-wise.*

This form, in terms of $\left\|(\hat{A}_1)_{ij} - (\hat{A}_2)_{ij}\right\|_1$, $\left\|(\hat{b}_1)_i - (\hat{b}_2)_i\right\|_1$, and $\|\hat{c}_1 - \hat{c}_2\|_1$ is particularly useful because the coefficients will be approximated in Fourier space. The reader familiar with energy estimates for elliptic equations will see that the requirements of the theorem are not the most general possible, and the proof can be modified to handle other cases individually when different estimates are desired.

Proof. Subtracting the first equation of (2.9) from the second, then adding and subtracting $A_1\nabla u_2$, $b_1\nabla u_2$, and $c_1 u_2$ gives

$$-\mathrm{div}(A_1\nabla w) - \mathrm{div}[(A_1 - A_2)\nabla u_2] + b_1 \cdot \nabla w + (b_1 - b_2)\nabla u_2 + c_1 w + (c_1 - c_2)u_2 = 0,$$

and after multiplying by w and integrating by parts, we glean

$$\begin{aligned}\lambda \int_\Omega |\nabla w|^2\, dx &\leq \int_\Omega \nabla w^T A_1 \nabla w + \left(c_1 - \frac{1}{2}\mathrm{div}(b_1)\right) w^2\, dx \\ &\leq \|A_1 - A_2\|_{op} \|\nabla u_2\|_2 \|\nabla w\|_2 + \|b_1 - b_2\|_\infty \|\nabla u_2\|_2 \|w\|_2 + \dots \\ &\quad \|c_1 - c_2\|_\infty \|u_2\|_2 \|w\|_2 .\end{aligned}$$

Using Poincare's inequality and $\|A_1 - A_2\|_{op} \leq d\,\|A_1 - A_2\|_\infty$,

$$\lambda \int_\Omega |\nabla w|^2\, dx \leq \left(d\,\|A_1 - A_2\|_\infty + C\,\|b_1 - b_2\|_\infty + C^2\,\|c_1 - c_2\|_\infty\right) \|\nabla u_2\|_2 \|\nabla w\|_2$$

and thus

$$\|\nabla w\|_2 \leq \lambda^{-1}\big(d\,\|A_1 - A_2\|_\infty + C\,\|b_1 - b_2\|_\infty + C^2\,\|c_1 - c_2\|_\infty\big)\,\|\nabla u_2\|_2\,. \quad (2.10)$$

Similarly multiplying the equation for u_2 by u_2, integrating by parts, and applying Poincare's inequality yields

$$\|\nabla u_2\|_2 \leq C\lambda^{-1}\,\|f\|_2\,.$$

Substituting this into (2.10) and using Poincare's Inequality again, we get

$$\|u_1 - u_2\|_{H^1} \leq C\lambda^{-2}\big(d\,\|A_1 - A_2\|_\infty + C\,\|b_1 - b_2\|_\infty + C^2\,\|c_1 - c_2\|_\infty\big)\,\|f\|_2\,.$$

The form stated in the theorem follows after

$$\|A\|_\infty = \max_{i,j}\|A_{ij}\|_\infty \leq \max_{i,j}\left\|\hat{A}_{ij}\right\|_1$$

and the analogous inequality with b. □

In practice memory is not a concern due to the convolutional structure of the matrix $\hat{L}_h$ representing an elliptic operator in Fourier space, but the sparse structure of the operator dramatically reduces computation complexity (Section 2.8.2).

2.5.3 Sparse Operator Approximation: Explicit Solver

The discrete analogue of Theorem 2.5.5 covers numerical schemes with implicit time steps, each of which require solving an elliptic problem with a sparsely approximated operator. Effectively, it allows us to estimate

$$\left\|Q^{-1} - P^{-1}\right\|_{op}$$

where P is a sparse matrix approximating that of the discretized full elliptic operator Q, and $\|\cdot\|_{op}$ refers to the L^2 matrix operator norm, or largest singular value. On the other hand, for explicit schemes, we are concerned about

$$\|Q - P\|_{op}$$

which we will consider directly.

Theorem 2.5.6. *Let L be the elliptic operator defined*

$$Lv = -\mathrm{div}(a\nabla v)$$

and let Q be its Fourier discretization

$$Q\hat{u} = k\,\hat{a}_h * (k\,\hat{u})$$

where k denotes the vector of Fourier mode frequencies and a_h is the discretized domain inhomogeneity coefficient in the elliptic operator. Analogously, let

$$P\hat{u} = k\,\hat{a}'_h * (k\,\hat{u}).$$

Then

$$\|Q - P\|_{op} \leq K^2\,\|\hat{a}_h - \hat{a}'_h\|_1$$

where K is the highest frequency on the grid.

In the case of a square grid $[1, \ldots, N]^d$, $K = N/2$. The result may be dismaying at first glance because it appears that the approximation error $\left\|\hat{a_h} - \hat{b_h}\right\|_1$ must be decreased faster than $O(1/N^2)$ just to remain stable. However, this type of bound is natural, since the operators' norms themselves are

$$\|P\|_{op} \approx \|Q\|_{op} = O(K^2).$$

The large operator norm is normalized by the stability condition $dt = O(dx^2)$, so one can think of these bounds in the update sense as:

$$\|(I - dtQ) - (I - dtP)\|_{op} \leq \|\hat{a}_h - \hat{a}'_h\|_1 .$$

Proof. The result is a simple consequence of Young's inequality: $\|f * g\|_2 \leq$

$\|f\|_1 \|g\|_2$. We have

$$\begin{aligned}
\|Q - P\|_{op} &= \sup_{||\hat{u}||=1} \|(Q-P)\hat{u}\|_2 \\
&= \|k(\hat{a}_h - \hat{a}'_h) * (k\hat{u})\|_2 \\
&\leq \|k(\hat{a}_h - \hat{a}'_h)\|_1 \|k\hat{u}\|_2 \\
&\leq K^2 \|\hat{a}_h - \hat{a}'_h\|_1 .
\end{aligned}$$

□

The proof clearly generalizes to hyperbolic operators of the form $Q\hat{u} = \hat{a} * (ik\,\hat{u})$ as well.

For an example, recall the forward Euler discretization of a parabolic PDE:

$$\hat{u}^{n+1} = (I - dt\hat{L}_h)\hat{u}^n,$$

over a time interval $[0, T]$. If $\|\hat{a}_h - \hat{a}'_h\|_1 = \delta$, approximating Q by P incurs an additional local truncation error of magnitude $\delta K^2 dt$ at each time step.

As the grid is refined, the CFL condition requires that $K^2 dt$ stay approximately constant, so that the approximation error per step remains approximately constant.

2.6 The Modified Equation Prespective

Using the variational principle for the explicit scheme applied to the parabolic equation yields the following first order optimality condition:

$$0 \in \hat{u}^{n+1} - (1 - dt\hat{L}_h)\hat{u}^n - dt\hat{f}_h + \mu\partial \left\|\hat{u}^{n+1}\right\|_1 ,$$

which is equivalent to

$$0 \in \frac{\hat{u}^{n+1} - \hat{u}^n}{dt} + \hat{L}_h\hat{u}^n - \hat{f}_h + \frac{\mu}{dt}\partial \left\|\hat{u}^{n+1}\right\|_1 .$$

Taking $\mu = \delta dt$ and formally sending dt and h to zero leads to

$$\hat{u}_t + \hat{L}\hat{u} - \hat{f} \in -\delta\partial \left\|\hat{u}\right\|_1, \tag{2.11}$$

or

$$\hat{u}_t + \hat{L}\hat{u} = \hat{f} - \delta p(\hat{u}) \tag{2.12}$$

where $p(\hat{u})$ denotes the particular element of the subdifferential so that the differential inclusion (2.11) is an equality. The sparse scheme applied to hyperbolic and elliptic problems yields analogous modified equations. We consider this to be the modified equation in the sense that the numerical scheme is directly solving this problem. The subgradient contribution is a vanishing 'compression' term which may be interpreted as a force which pushes the solution u toward the nearest (in the L^1 proximal sense) union of low dimensional subspaces spanned by the Fourier basis.

Well-posedness for the modified equation is guaranteed via the theory of differential inclusions on Banach spaces (*e.g.* [9, 26]). The theorem below summarizes these results in the current context.

Theorem 2.6.1 (Well-posedness)**.** *Let $u(t)$ satisfy the differential inclusion*

$$\partial_t u(t) \in -A(u(t)) - \delta\partial||\hat{u}(t)||_{L^1}$$

with $u(0)$ in the domain of the monotone (single-valued) operator A. Then for all $\delta \geq 0$, there exists a unique solution $u(t)$ defined for all $t \geq 0$ which is the solution to

$$\partial_t u(t) = -A(u(t)) - \delta p(\hat{u}(t))$$

for some $p \in \partial||\hat{u}(t)||_{L^1}$.

Lastly, we mention that if we want to directly compare the error between the solutions of the original and modified equations, the error grows linearly in time (at worst).

Theorem 2.6.2. *Let u be the solution to*

$$u_t + Lu = f$$

and let u_δ solve

$$(\hat{u}_\delta)_t + \hat{L}\hat{u}_\delta - \hat{f} \in \delta\partial \left\|\hat{u}_\delta\right\|_1 .$$

Then

$$||u(t,\cdot) - u_\delta(t,\cdot)||_2 \leq 2\delta t.$$

The proof is direct and can be found in the Appendix. Similar results are easily proved for the elliptic and hyperbolic cases using only that $\left\|p(\hat{u})\right\|_\infty \leq 1$ and standard energy estimates; this approach also provides a simple alternate proof.

2.7 Denoising Perspective

Soft thresholding also appears in early methods for signal denoising using wavelets [30]. We refer the reader to that work for full details, and list here only the analogues of its major results in the current context.

Consider the following denoising problem: we wish to recover a signal $f \in \mathbb{R}^n$ from noisy observations $d = f + w$, $||w||_1 \leq \mu$, by soft-thresholding DFT coefficients by μ. This approach enjoys the following properties:

- (Smoothing) The recovered signal f_μ satisfies $||f_\mu||_{H^k} \leq ||f||_{H^k}$ for any Sobolev norm $\left\|\cdot\right\|_{H^k}$. In particular, $|\hat{f}_\mu(k)| \leq |\hat{f}(k)|$ for all frequencies k.
- (Near optimality) f_μ is near-minimax:

  $$\sup_{||f||_{H^k} \leq C_1} \sup_{||w||_1 \leq \mu} ||f_\mu - f||_{l_n^2}^2 \leq 4 \inf_{\tilde{f}} \sup_{||f||_{H^k} \leq C_1} \sup_{||w||_1 \leq \mu} ||\tilde{f}(d) - f||_{l_n^2}^2$$

  where $\tilde{f}(d)$ is any other estimator of f.

The smoothing property guarantees that the recovered signal is 'noise-free'; the near optimality property guarantees that for worst-case signals of bounded Sobolev norm and noise of bounded ℓ^1 norm, the result recovered by soft thresholding is nearly the 'optimal' (see [30]).

Next, consider the solution u^ϵ to the standard parabolic multiscale problem

$$\frac{\partial u^\epsilon}{\partial t} - \frac{\partial}{\partial x}\left(a(x, x/\epsilon)\frac{\partial u^\epsilon}{\partial x}\right) = 0 \quad \text{on } [0, 2\pi] \text{ periodic}, \quad u^\epsilon(x, 0) = u_0^\epsilon(x).$$

The theory of asymptotic homogenization (*e.g.* [70]) can be used to show that at time point t^n, the exact solution u^ϵ satisfies

$$u^\epsilon(x, t^n) = u_0(x, t^n) + \epsilon u_1(x, x/\epsilon, t^n) + \epsilon^2 R(x, t^n)$$

with $|\hat{R}(x, t)| \leq C$. This expansion is valid as long as we assume that the equation is taken on a periodic domain and $a(x)$ is as smooth as we like. For a numerical solution, the asymptotic expansion can be easily modified to include truncation error τ^{n+1} as follows: if we let v^{n+1} denote the numerical solution at time t^{n+1}, then

$$v^{n+1} = u_0(x, t^{n+1}) + \epsilon u_1(x, x/\epsilon, t^{n+1}) + \epsilon^2 R(x, t^{n+1}) - \tau^{n+1}.$$

This form allows us to draw a connection between the denoising and homogenization problems: for an appropriate threshold choice μ, the compressive spectral method denoises v^{n+1} as

$$v^{n+1} = \underbrace{u_0(x, t^{n+1}) + \epsilon u_1(x, x/\epsilon, t^{n+1})}_{\text{signal}} + \underbrace{\epsilon^2 R(x, t^{n+1}) - \tau^{n+1}}_{\text{noise}}$$

and attempts to recover the first terms in the asymptotic expansion. These interpretation is valid between any two time steps, but may not hold globally.

2.8 Efficient Implementation

In this section, we describe important details pertaining to the numerical method and algorithm considerations. Using a concrete example, we show that a favorable

complexity can be achieved.

2.8.1 The Proximal-Galerkin Algorithm

The implicit scheme described above requires fast minimization of the energy (2.8), and differs from many cases where L^1 regularization is added because the problem, *e.g.* compressed sensing [17], TV minimization [75], or basis pursuit [21], is ill-posed without it. For the multiscale PDE problem, this is not the case since an appropriately discretized version of (2.8) will be well-posed and can be solved by inverting a linear system

$$Q\hat{u} = \hat{f}$$

where Q is a positive-definite (and even sparse, in physical rather than Fourier space) matrix. If the elliptic operator is discretized appropriately, fast and extensively studied preconditioned conjugate gradient solvers are available. So, to be competitive, the compressive implicit scheme must leverage sparsity of the solution $\hat{u}$ to perform the (approximate) linear inversion $Q\hat{u} = \hat{f}$ quickly. For this purpose, we propose the hybrid proximal gradient descent and Galerkin approximation algorithm described below, which is related to the procedure described in [28].

First, let D be the diagonal part of Q. Since Q is the matrix corresponding to a Fourier-space discretized elliptic operator, D is the matrix corresponding to a multiple the Fourier-space discretized Laplacian. We take $n \sim 10$, $\mu > 0$, $\omega > 0$, and initialize the solution to be zero (*i.e.* $\hat{u} = 0$).

The Proximal-Galerkin Algorithm

for $j = 1{:}n$ **do**

$\hat{u} = \text{shrink}\big(\hat{u} + \omega D^{-1}(\hat{f} - Q\hat{u}), \mu\big)$;

end for

set $I = \text{supp}(\hat{u})$;

set $\hat{u} = \arg\min w : \ \text{supp}(w) \subseteq I \frac{1}{2} \left\| Qw - \hat{f} \right\|_2^2$;

Return $\hat{u}$.

The algorithm begins with a few iterations of the proximal gradient method applied to the energy

$$E(w) = \mu \left\| w' \right\|_1 + \frac{1}{2} w'^T Q' w' - w'^T \hat{f}'$$

where

$$Q' = D^{-1/2} Q D^{-1/2},$$
$$\hat{f}' = D^{-1/2} \hat{f},$$
$$w' = D^{1/2} w.$$

This is a simple Jacobi preconditioning of the analogous energy with Q, w, and $\hat{f}$. Rather than iterating proximal gradient to convergence, which would be too slow, the algorithm stops after just a few iterations with rough approximation. The support of that solution is used to identify the Fourier modes with largest magnitude coefficients, and then a Galerkin approximation is computed over those modes. Due to sparsity in the Fourier domain, the linear solve associated with the Galerkin part is small and inexpensive– computational complexity depends on the grid mesh size only through the sparsity of the solution.

2.8.2 Algorithm Complexity

The pseudospectral approach of computing the convolution

$$k\,\hat{a} * (k\,\hat{u})$$

uses an FFT, and for an N-gridpoint problem this reduces the computational complexity per iteration from $O(N^2)$ to $O(N \log N)$. We now consider the computational complexity of the sparse spectral method, which must be comparable to $O(N \log N)$ to be practical.

Suppose that the sparsely approximated operator is defined $P\hat{u} = k\,\hat{a}' * (k\,\hat{u})$, where the sparsity (number of nonzeros) of $\hat{a}'$ is m, and that the sparsity of $\hat{u}$ is r. By treating the $\hat{a}' * \hat{u}$ sparse convolution as a summation of sparse vectors, it can be accomplished with complexity

$$O(mr \min(\log r, \log m)), \tag{2.13}$$

free of any dependence on the full problem size N, by storing the sparse vectors $\hat{a}'$ and $\hat{u}$ as sorted linked lists and computing the sum as a merge operation, with a priority queue. For the modest one-time cost of initializing a length N array, the complexity can be decreased to

$$O(mr) \tag{2.14}$$

by leaving the sparse vectors unsorted. We iterate over the mr nonzero coefficients which must be added, and use an auxiliary array keep track of the partial result. When a new coefficient of the partial result becomes nonzero, it is placed in a growing list of indices. After we have visited each of the mr coefficients to be added, we iterate over the list of nonzero indices, perform the shrink operation on the corresponding auxiliary array entry holding the partial result, and copy the outcome into a list which holds the final result. Along the way, we 'zero out' the entry of the partial result array, never incurring another $O(N)$ cost.

Finally, if the problem is elliptic or requires implicit time steps and the Proximal-Galerkin algorithm is used, the complexity includes a term

$$O(r^3),$$

the cost of the Galerkin linear solve over the support found with proximal gradient.

Both (2.13) and (2.14) are preferable to the $O(N \log N)$ cost of the pseudospectral method for very sparse problems and in the homogenization limit discussed next in Section 2.8.3. For the numerical examples considered in this limit, m and r stay approximately constant, leading to computation time which does not increase as the grid is refined.

One key to the effective application of the sparse spectral method is proper discretization. For a typical homogenization problem, we are interested in the solution of an equation such as

$$-\text{div}(a(x/\epsilon)\nabla u) = f$$

for ϵ close to zero, and we might choose the inhomogeneity coefficient

$$a(x) = 1 + \frac{1}{2}\sin \pi x.$$

This choice is ideally sparse in the Fourier domain, with only three nonzero entires regardless of N, using the standard uniform grid. If $\epsilon = 1/1000$, then $\hat{a}$ still has only three nonzeros. However, choosing $\epsilon = \frac{1}{707\sqrt{2}}$ results in $\hat{a}$ being completely dense. These two choices of ϵ differ by less than 10^{-6}, and the first leverages extreme sparsity in the problem while the second does not. This example shows that it is prudent to assume a certain relationship between the grid spacing and ϵ, considered next.

2.8.3 Homogenization Limit

For homogenization problems in particular, where one is interested in the limit $\epsilon \to 0$, we can keep $N\epsilon$ fixed as the grid is refined. Empirically, we have observed that this keeps the sparsity of the operator and of the solution approximately constant. For a simple case of this $N\epsilon = c$ (c a constant) limit, the following theorem guarantees the sparsity of the operator remains fixed along a subsequence.

Theorem 2.8.1. *Let L_ϵ be the elliptic operator defined*

$$L_\epsilon v = -\mathrm{div}(a(x/\epsilon)\nabla v),$$

and let

$$Q_{\epsilon,N} u = k\, \hat{a_N} * (k\, \hat{u})$$

be its Fourier discretization on an $N-$point discretization of $[0, 2\pi)$. Then $Q_{\epsilon,N}$ and $Q_{\epsilon/2,2N}$ are equally sparse: that is,

$$\#\{k : |\hat{a_{2N}}(k)| \geq \lambda\} = \#\{k : |\hat{a_N}(k)| \geq \lambda/2\} \tag{2.15}$$

for all $\lambda > 0$.

See the appendix for a proof. Note that the theorem assumes the standard definition of the DFT on N grid points,

$$\mathcal{F}_N[a(x)](k) = \sum_{j=0}^{N-1} a\left(\frac{2\pi j}{N}\right) e^{-2\pi i j k/N},$$

which is not unitary. This accounts for the appearance of $\lambda/2$ rather than λ on the right hand side of (2.15). This factor cancels out in the end because with this definition of the DFT, the ℓ^1 norm in Theorems 2.5.5 and 2.5.6 should be scaled by $1/N$.

The complexities (2.13) and (2.14) become very favorable in the $N\epsilon = c$ limit, where m and r remain nearly constant or grow approximately logarithmically with N as the grid is refined. In each case we observed, the overall algorithm complexity is linear or sub-linear in N.

2.9 Numerical Examples

In [76], the authors demonstrated the effective application of the compressive spectral method to a variety of problems. Here, we expand on those results and give examples of the additions to the method proposed in this chapter: the implicit scheme and sparse operator approximation.

2.9.1 Transport Equation, 1D

The PDE considered is the traveling wave equation:

$$u_t + a(x)u_x = 0,$$
$$x \in [0, 2\pi] \text{ periodic},$$
$$u(x, 0) = \sin(x)$$

with oscillatory coefficient

$$a(x) = \frac{1}{8} \exp\left(\frac{0.6 + 0.2\cos x}{1 + 0.7 \sin 128x}\right).$$

The update is given by leap frog time discretization:

$$\hat{u}^{n+1} = \hat{u}^{n-1} - 2dt\hat{a} * (ik\,\hat{u}^n).$$

We choose the above form for a throughout this section, because it is less sparse than simple trigonometric functions.

The grid sizes considered are $N = 2^{10}, \ldots, 2^{14}$ and the values of other parameters are $dt = 6.25 \times 10^{-6}$, $\|\hat{a} - \hat{a}'\|_1 = 10^{-2}$, $\mu = 1.2 \times 10^{-5}$, and the simulation is run to a final time $t = 0.5$.

Figure 2.3 shows the full spectral and compressive spectral (sparse operator/sparse solution) solutions on coarse and fine scales. The compressive scheme correctly captures the largest Fourier coefficients of the solution, discarding all but 3.7%, and the operator approximation discards all but 2.6%. The “true" solution was computed on a fine grid with finite difference methods.

Figure 2.4 shows the L^2 error and sparsity of the compressive spectral approximations as the grid is refined with dt held constant. Error is computed as the L^2 distance to the full spectral solution. The error of the sparse operator/sparse solution scheme is dominated by the sparse approximation of the solution; spurious modes in the leap frog scheme make a sparse approximation of it difficult. Over the range of grids considered, sparsity of the operator eventually becomes

constant while sparsity of the solution grows about linearly. The complexity of the compressive spectral method is thus linear in N over the grid sizes considered.

Figure 2.5 considers the same problem but with a resonant forcing term

$$f(x) = e^{\sin(x/128)^2}$$

with $N = 2048$ and all other parameters the same as the non-forced problem. The solution has 11.3% nonzero Fourier coefficients, with $\|u_{\text{full}} - u_{\text{sparse}}\|_2 = 2.5\times10^{-3}$. The resonant forcing causes sharp and irregular oscillations at the fine scale, which make the problem less sparse, but the compressive scheme still captures the correct behavior.

2.9.2 Elliptic Problem, 1D

The PDE considered is the elliptic problem:

$$-(a(x)u_x)_x = \sin 2x,$$
$$x \in [0, 2\pi] \text{ periodic},$$
$$\int u\,dx = 0$$

with

$$a(x) = \exp\left(\frac{0.6 + 0.2\cos x}{1 + 0.7\sin x/\epsilon}\right)$$

such that $N\epsilon = 8$, and usual spectral operator discretization:

$$\hat{L}_h\hat{u} = k\,\hat{a}_h * (k\,\hat{u}) = \hat{f}.$$

This time we consider the homogenization limit, keeping $N\epsilon = 8$ with $\epsilon = \frac{1}{64}, \frac{1}{128}, \ldots, \frac{1}{1024}$, and set $\|\hat{a} - \hat{a}'\|_1 = 1\times10^{-4}$. Parameter values for the Proximal-Galerkin algorithm are $n = 10$, $\mu = 5\times10^{-8}$, and $\omega = 5\times10^{-3}$.

Figure 2.6 shows the full spectral and compressive spectral solutions on coarse and fine scales. Both the sparse solution and operator approximation keep 8.5%

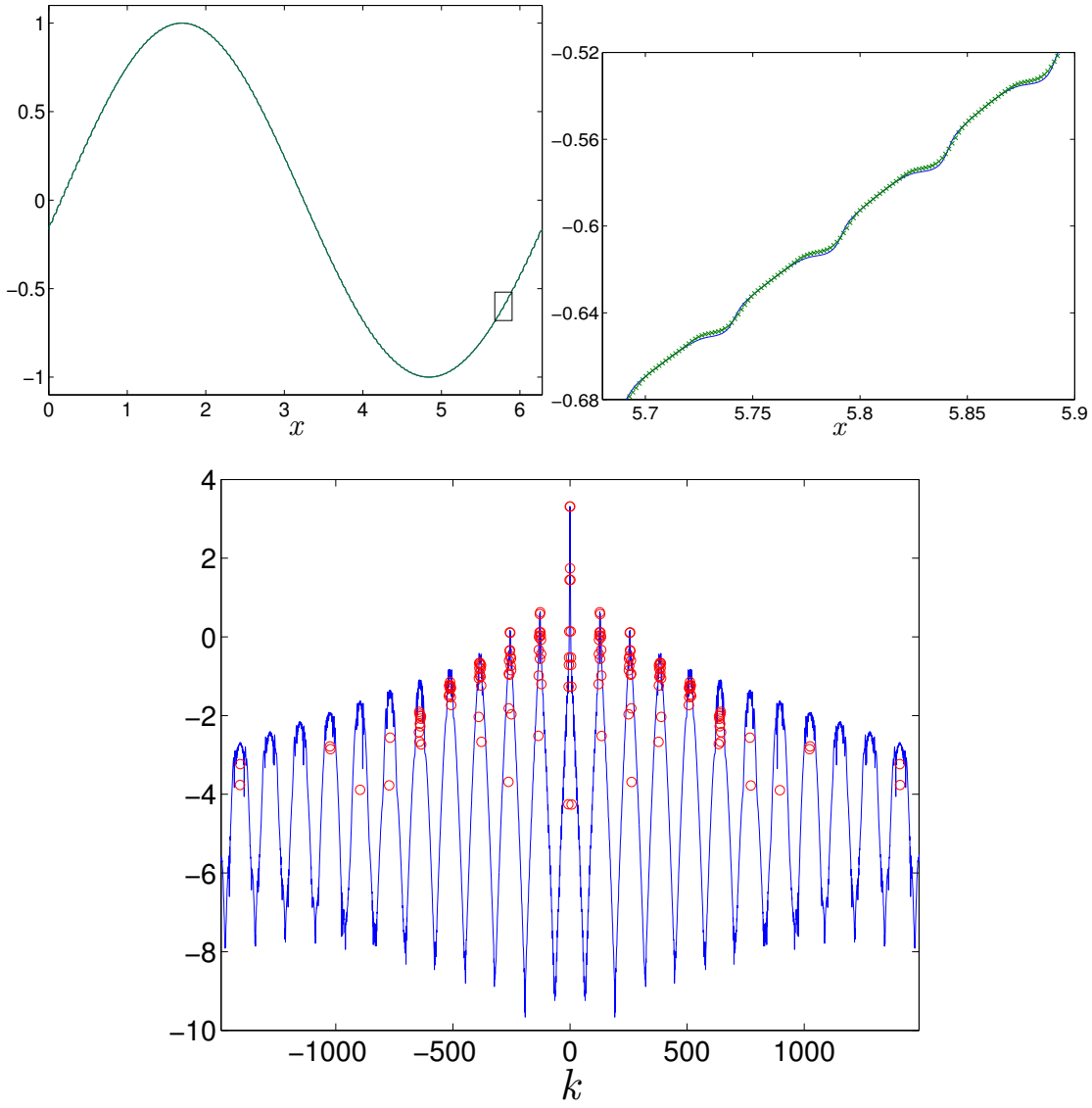


Figure 2.3: **Left:** True (blue) and sparse operator/sparse solution (green) solutions in physical space. The two curves lie almost on top of each other. **Right:** Zoomed in true (blue) and sparse (green '×') solutions. **Bottom:** True (blue) and sparse (red '∘') solutions in Fourier space. $N = 4096$, operator nonzeros = 107, solution nonzeros = 153.

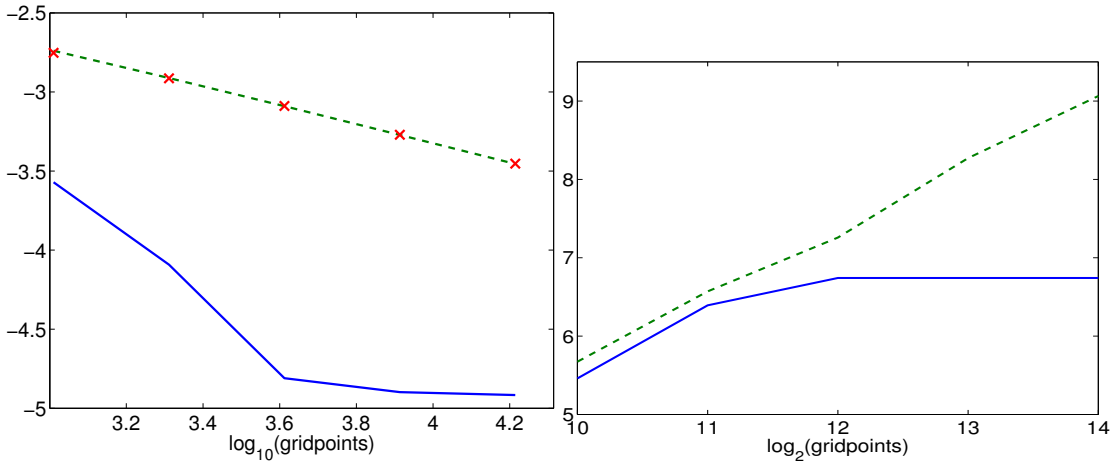


Figure 2.4: **Left:** Sparse operator/full solution (blue), full operator/sparse solution (green, dashed), and sparse operator/sparse solution (red $\times$) L^2 distance to the full spectral solution as the grid is refined. The y axis has a $\log_{10}$ scale. **Right:** Number of nonzero Fourier coefficients of the operator (blue) and solution (green, dashed) as the grid is refined. The y axis has a $\log_2$ scale.

of the coefficients. Note that the full result and sparse operator/sparse solution result lie almost on top of each other, even at the resolution of the fine scale.

Figure 2.7 shows error (L^2 distance to the full spectral solution) and sparsity under refinement. "Sparse operator" refers to the solution obtained with the sparsely approximated operator, using either a high accuracy conjugate gradient solve or the Proximal-Galerkin algorithm. "Sparse solution" refers to the use of the Proximal-Galerkin algorithm, with either the full or sparse operator.

Approximation error does not increase while both solution and operator sparsity remain approximately constant, leading to computation time approximately independent of N. With $N = 2^{13}$, the sparse approximation maintains six digits of accuracy with only 1.1% of the coefficients of both the operator and the solution.

Figure 2.8 illustrates that for a fixed number of nonzero coefficients, the sparse operator approximation incurs smaller error than the solution approximation.

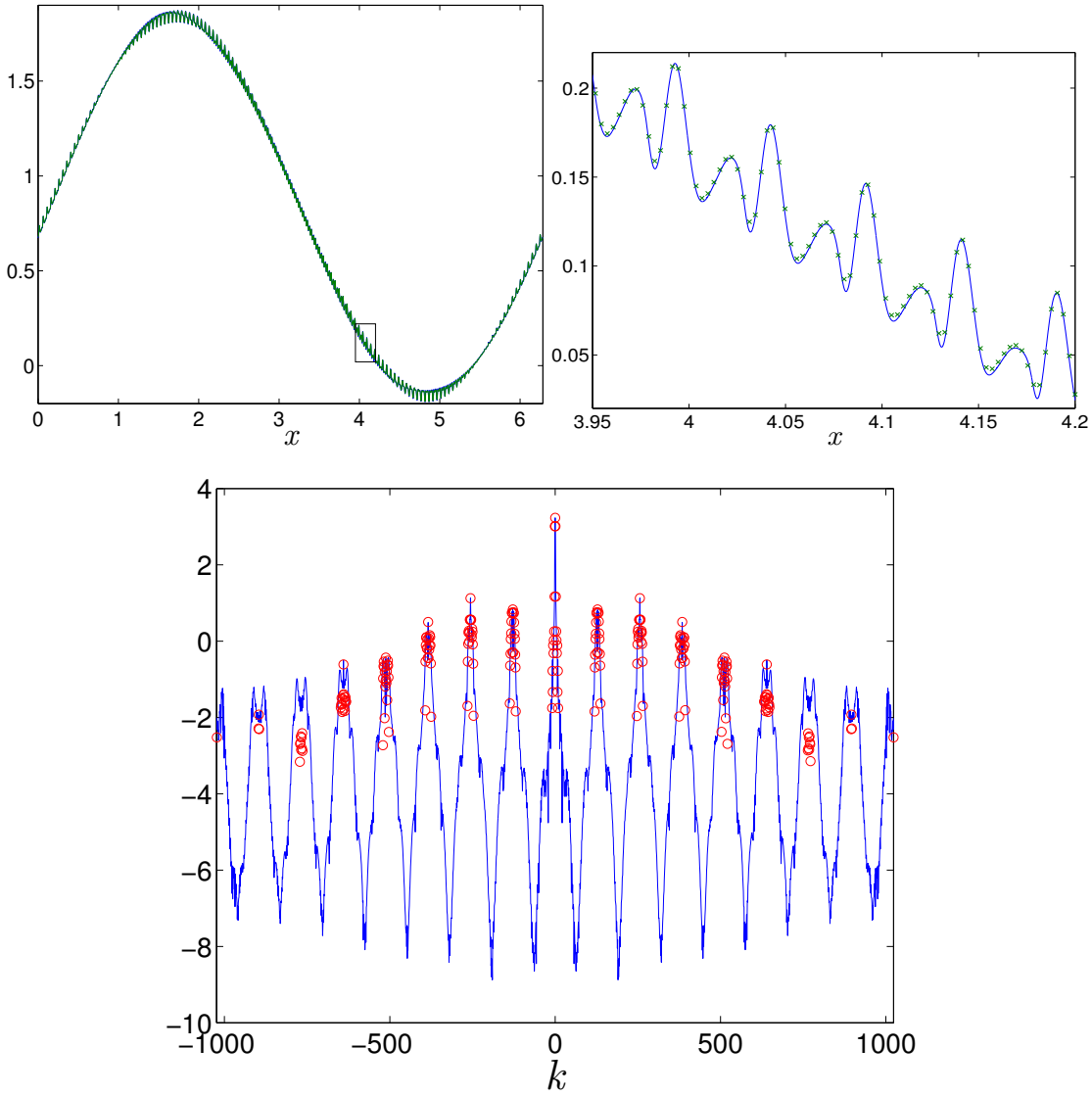


Figure 2.5: **Left:** True (blue) and sparse operator/sparse solution (green) solutions with resonant forcing term in physical space. **Right:** Zoomed in true (blue) and sparse (green '×') solutions. **Bottom:** True (blue) and sparse (red '○') solutions in Fourier space. $N = 2048$, operator nonzeros = 86, solution nonzeros = 231.

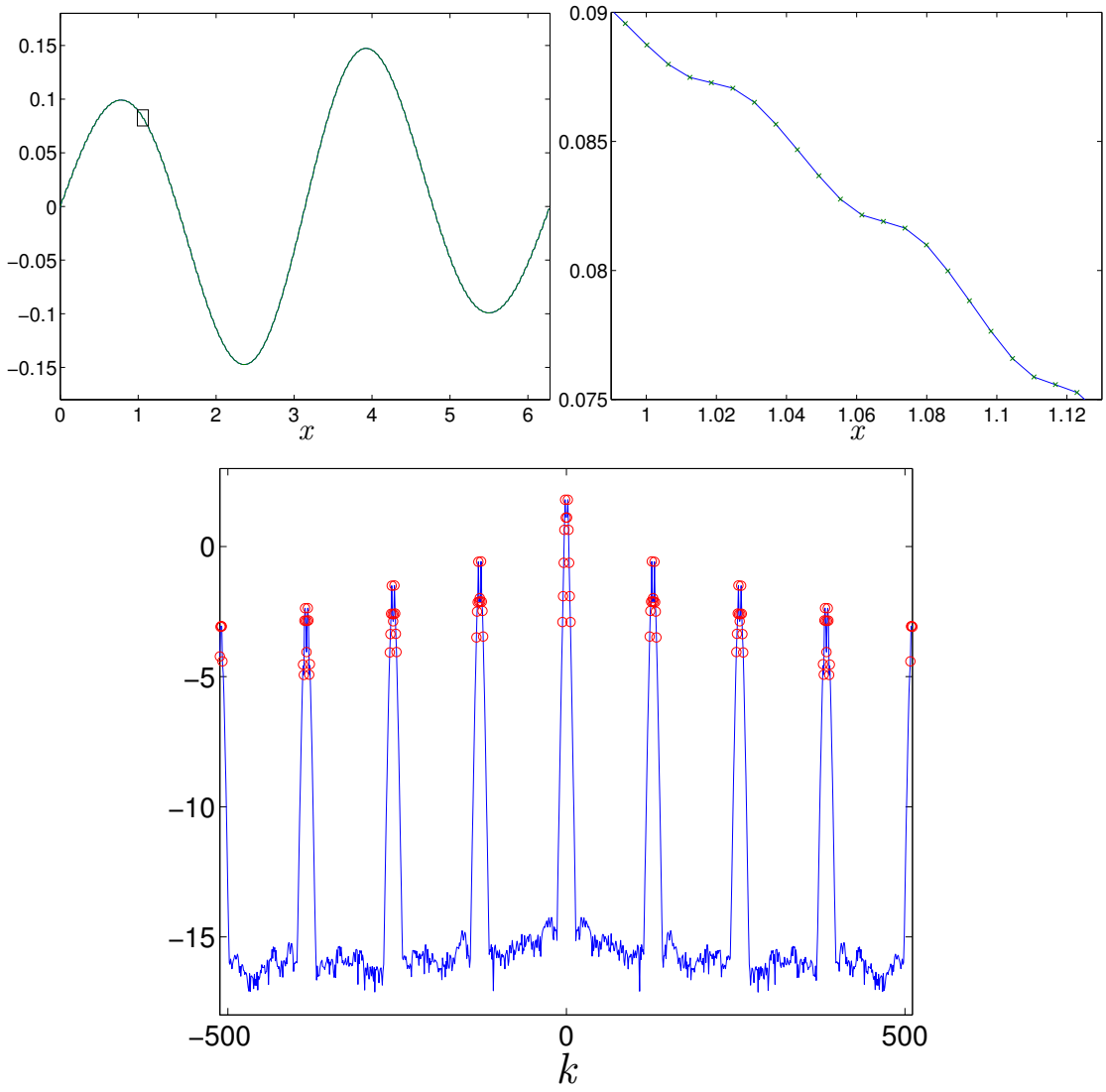


Figure 2.6: **Left:** True (blue) and sparse operator/sparse solution (green) solutions in physical space. The small rectangle shows the axis limits of the zoomed in plot to the right. **Right:** Zoomed in true (blue) and sparse (green '×') solutions. **Bottom:** True (blue) and sparse (red '∘') solutions in Fourier space. $N = 1024$, operator nonzeros = 86, solution nonzeros = 87.

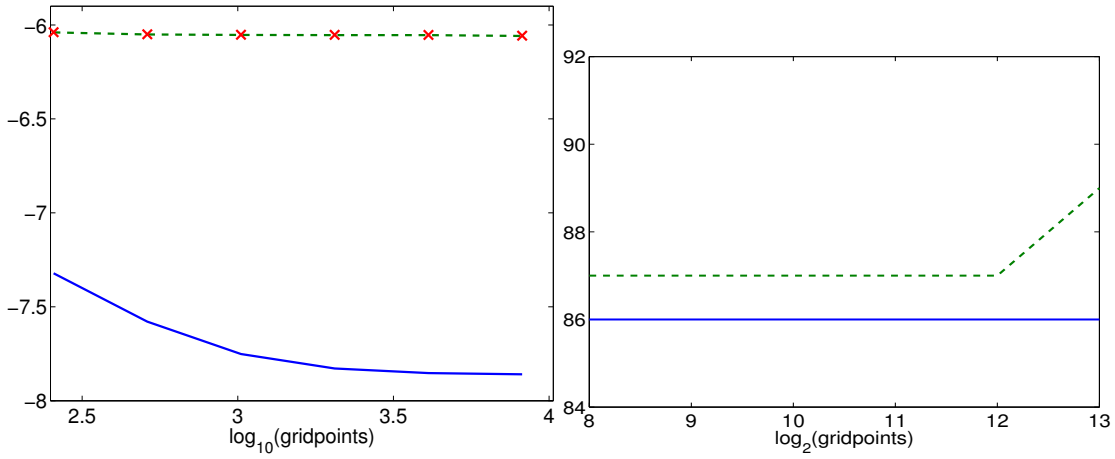


Figure 2.7: **Left:** Sparse operator/full solution (blue), full operator/sparse solution (green, dashed), and sparse operator/sparse solution (red $\times$) error under the homogenization limit. The y axis has a $\log_{10}$ scale. **Right:** Number of nonzero Fourier coefficients of the operator (blue) and solution (green, dashed) as the grid is refined.

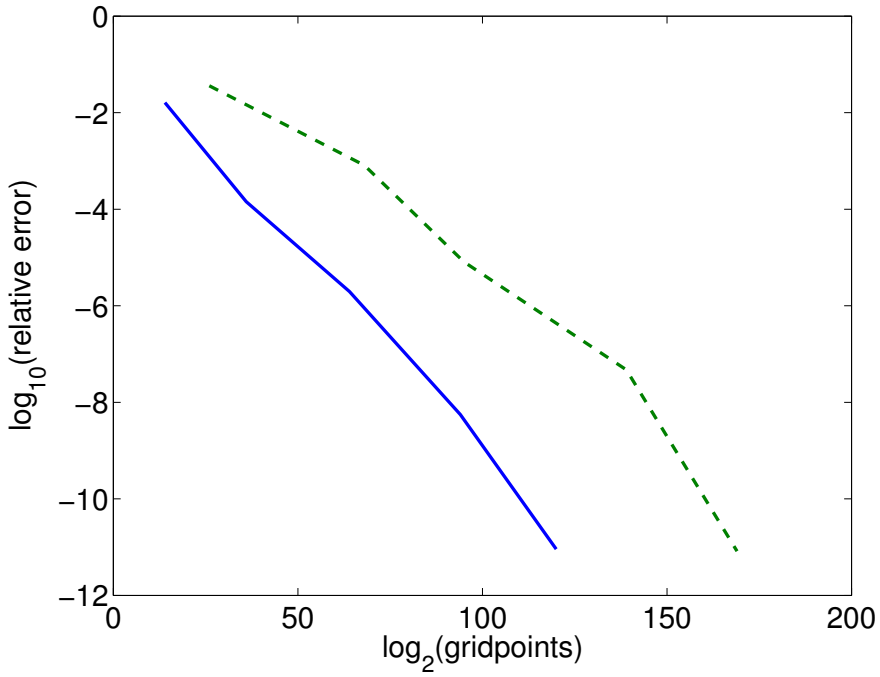


Figure 2.8: Pareto curves showing the tradeoff between approximation error and sparsity of the operator (blue) and solution (green, dashed).

2.9.3 Parabolic Problem, 1D

The PDE we consider here is the parabolic equation:

$$u_t - (a(x)u_x)_x = 0,$$
$$x \in [0, 2\pi] \text{ periodic},$$
$$u(x, 0) = 1 + \cos(x - \pi)$$

with
$$a(x) = \exp\left(\frac{0.6 + 0.2\cos x}{1 + 0.7\sin x/\epsilon}\right)$$

We again consider the $N\epsilon = 8$ limit, $\epsilon = \frac{1}{64}, \frac{1}{128}, \ldots, \frac{1}{1024}$, and set $\|\hat{a} - \hat{a}'\|_1 = 1\times10^{-2}$ and $dt = 1\times10^{-2}$ for all N. Parameter values for the Proximal-Galerkin algorithm are $n = 10$, μ ranges from 5×10^{-6} to 6.4×10^{-6}, and $\omega = 1\times10^{-2}$.

Figure 2.9 compares the solutions on coarse and fine scales. The sparse solution retains 3.2% of the coefficients and the operator is also approximated with 3.2%. Figure 2.10 shows error and sparsity under refinement. Approximation error decreases while sparsity of both operator and solution stay constant. The overall complexity is thus constant in N over the range of grid sizes considered. For this problem, sparse approximation of the operator incurs most of the error.

2.9.4 Elliptic Problem, 2D

We consider the elliptic problem

$$-\text{div}(a(x)\nabla u) = 10\sin x \sin y,$$
$$x, y \in [0, 2\pi] \text{ periodic},$$
$$\int u\, dxdy = 0$$

with
$$a(x, y) = \exp\left(\frac{0.6 + 0.2\cos x}{1 + 0.7\sin x/\epsilon} + \frac{0.6 + 0.2\cos y}{1 + 0.7\sin y/\epsilon}\right)$$

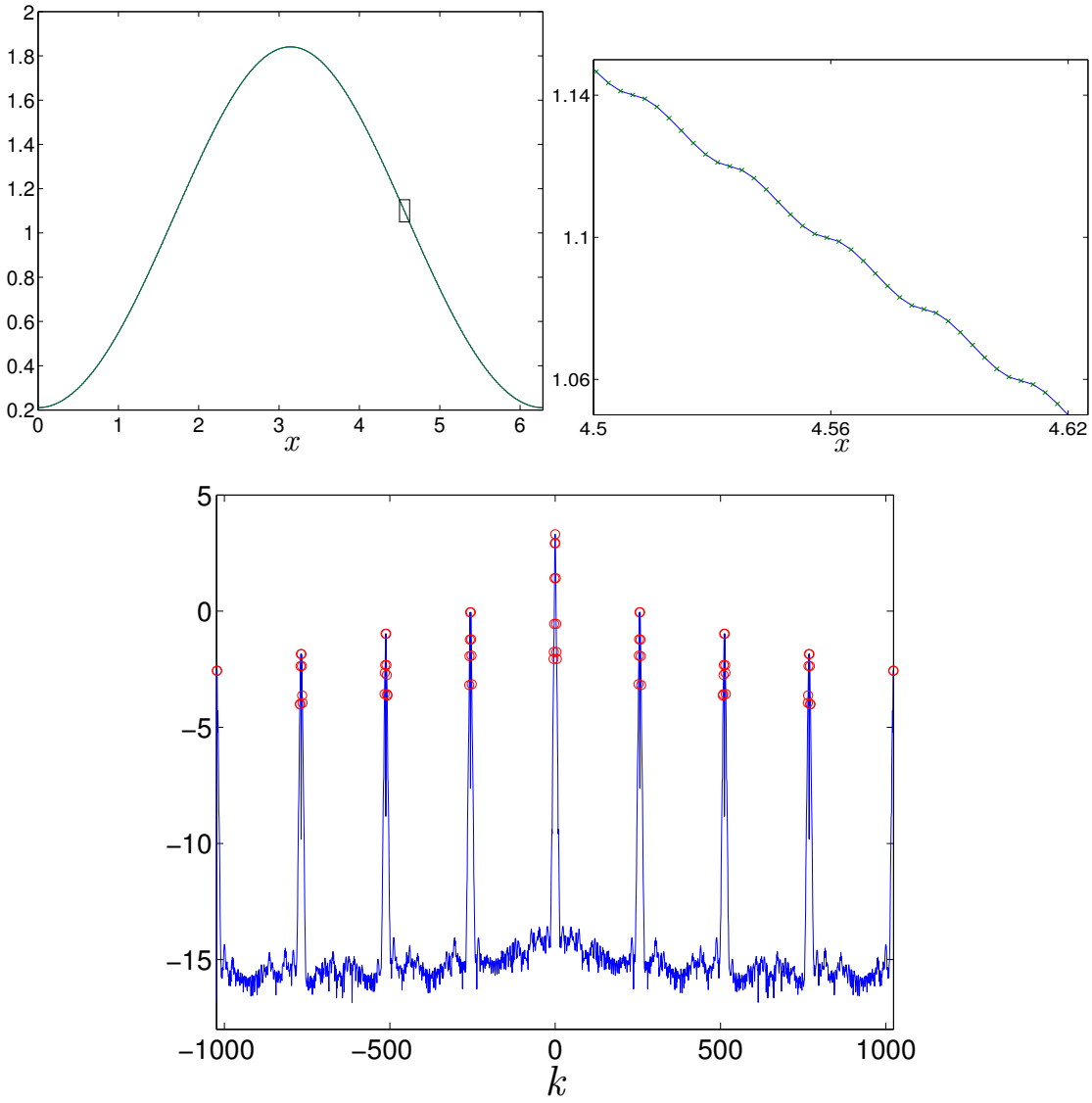


Figure 2.9: **Left:** True (blue) and sparse operator/sparse solution (green) solutions in physical space. **Right:** Zoomed in true (blue) and sparse (green '×') solutions. **Bottom:** True (blue) and sparse (red '∘') solutions in Fourier space. $N = 2048$, operator nonzeros $= 64$, solution nonzeros $= 65$.

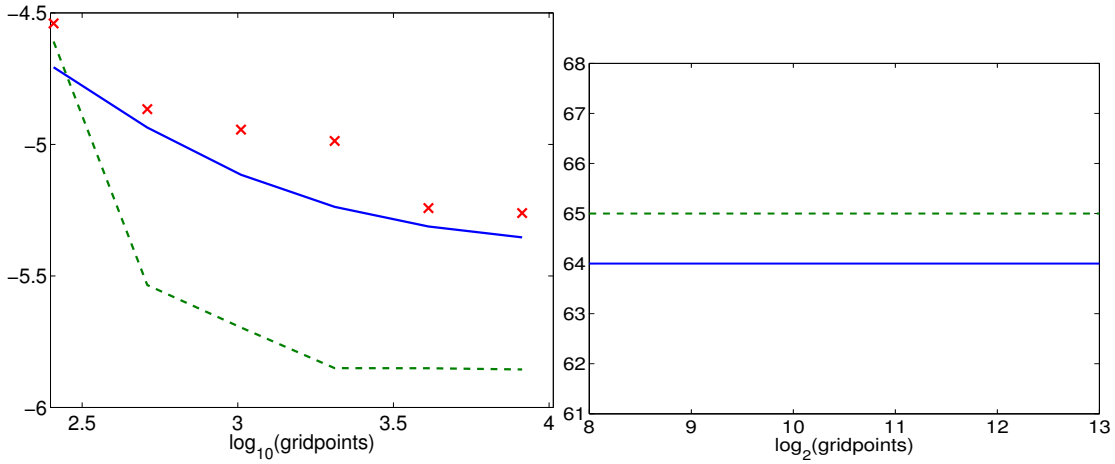


Figure 2.10: **Left:** Approximation error of the sparse operator/full solution (blue), full operator/sparse solution (green, dashed), and sparse operator/sparse solution (red $\times$) error under the homogenization limit. The y axis has a $\log_{10}$ scale. **Right:** Number of nonzero Fourier coefficients of the operator (blue) and solution (green, dashed) are constant as the grid is refined.

on an $N \times N$ grid such that $N\epsilon = 8$, with $\epsilon = \frac{1}{16}, \frac{1}{32}, \ldots, \frac{1}{256}$ and $\|\hat{a} - \hat{a}'\|_1 = 1$. Parameter values for the Proximal-Galerkin algorithm are $n = 20$, μ between 4×10^{-4} and 32×10^{-4}, and $\omega = 2\times10^{-2}$.

Because the full and spectral solutions are very close to each other in physical space and an overlaid comparison of surfaces is difficult, Figure 2.11 shows the solutions on a log scale in Fourier space. Of the 2^{20} coefficients in the full solution, the sparse solution and operator retain just 0.2% while maintaining four digits of accuracy. Figure 2.12 shows that approximation error decreases slightly with constant sparsity and computation time. For some grid sizes, the sparse operator/sparse solution scheme actually attains a lower error than the sparse operator/full solution scheme, evidence of the denoising effect discussed in Section 2.7.

To compare the Fourier coefficients of the full and sparse solutions more ac-

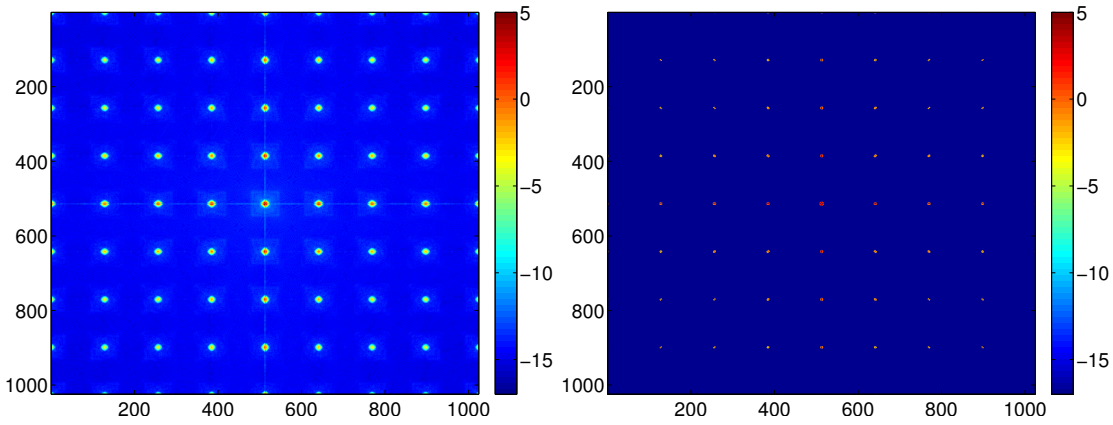


Figure 2.11: Full (left) and sparse (right) solutions on a log scale in Fourier space. Note that the great majority of coefficients in the sparse solution are exactly zero. $N = 1024$, $\epsilon = \frac{1}{128}$, operator nonzeros $= 1972$, solution nonzeros $= 1874$.

curately, the left panel of Figure 2.13 shows the magnitude of the 4500 largest Fourier coefficients of the true solution sorted in descending order. The magnitude of the corresponding sparse solution Fourier coefficients is also shown, with an upward bias to account for all the wave numbers not present. The right panel shows the fraction of full solution wave numbers which are captured by the sparse scheme. The compressive scheme correctly identifies all 500 of the largest modes in the full solution, and about 68% of the full solution's largest 1800 modes.

2.10 Conclusion

In this chapter, we have proposed a sparse operator approximation and an efficient method for extending the work of [76] to implicit solvers (Section 2.4). We have proven the convergence of the original compressive spectral scheme [76] and the new variants, including a modified equation which shows the effect of soft thresholding is equivalent to including an L^1 subgradient term in the PDE. Also, we connect the homogenization problem with that of signal denoising via wavelet thresholding. For PDE with sparse initial data or forcing terms, the new methods

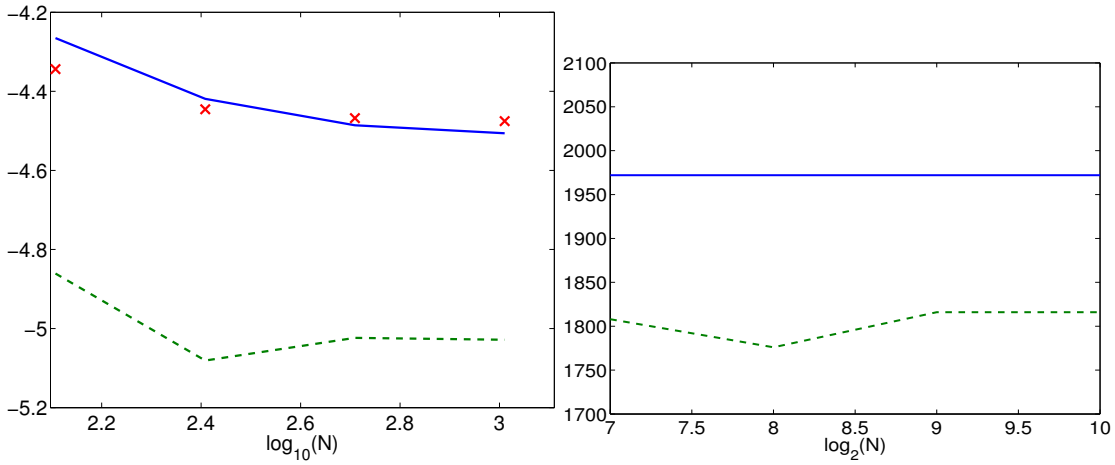


Figure 2.12: **Left:** Approximation error of the sparse operator/full solution (blue), full operator/sparse solution (green, dashed), and sparse operator/sparse solution (red $\times$) error under the homogenization limit. The y axis has a $\log_{10}$ scale. **Right:** Number of nonzero Fourier coefficients of the operator (blue) and solution (green, dashed) are constant as the grid is refined.

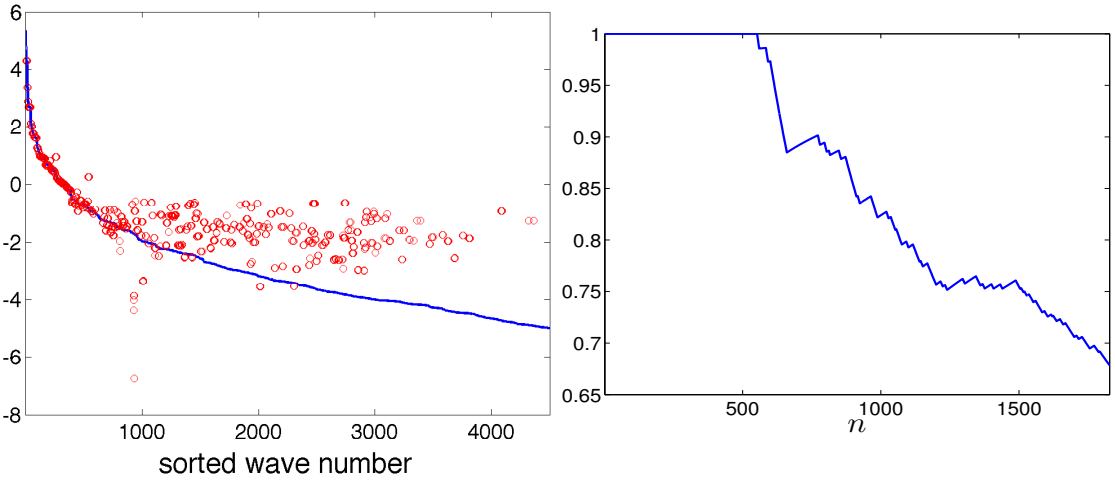


Figure 2.13: **Left:** energy spectrum decay of the full and sparse solutions. The plot shows just the largest 4500 coefficients of the full solution, the support of which contains all coefficients of the sparse solution. **Right:** fraction of sparse modes appearing among the largest n true modes, as a function of n.

are asymptotically preferable to the pseudospectral approach. The methodology presented here could be translated to other psuedospectral methods which employ alternative bases. Computationally, this amounts to replacing the Fast Fourier transforms in the psuedo-codes above with the appropriate transformation. This could be useful in cases where the solutions are sparse against another known basis.

2.11 Appendix

Before giving the proofs of the theorems from Section 2.5.1, we recall the definition of Bregman Distance (also known as Bregman Divergence).

Definition 2.11.1. Let J be a convex function and u, v be points in the domain of J. Also let p be an element of the subdifferential of J, *i.e.* $p \in \partial J(v)$. We define the *Bregman Distance* between u and v as

$$D_J^p(u, v) = J(u) - J(v) - \langle p, u - v \rangle.$$

In general, the Bregman Distance is not symmetric and does not obey the triangle inequality, so it is not a distance in the typical sense.

In what follows, we will also use basic facts regarding monotone operators.

Definition 2.11.2. Let A be a multi-valued map from V into itself. We call A *monotone* if and only if for any $u, v \in Dom(A)$ and any values Au and Av might take on,

$$\langle u - v, Au - Av \rangle \geq 0.$$

If $A = \partial F$ is the subdifferential of a convex function, then it is monotone.

We can now give the proof of theorem 2.5.1, in which we omit hats for notational clarity.

Proof. Consider the iterations for arbitrary points u^n and v^n:

$$\mu p(u^{n+1}) + \frac{u^{n+1} - u^n}{dt} = -\hat{L}_h u^n + f$$
$$\mu p(v^{n+1}) + \frac{v^{n+1} - v^n}{dt} = -\hat{L}_h v^n + f.$$

By taking the difference between these two equations we arrive at

$$\mu(p(u^{n+1}) - p(v^{n+1})) + \frac{1}{dt}(u^{n+1} - v^{n+1}) - \frac{1}{dt}(u^n - v^n) = -\hat{L}_h(u^n - v^n)$$

and taking the inner product of this equation with $u^{n+1} - v^{n+1}$ yields

$$\mu \left\langle p(u^{n+1}) - p(v^{n+1}), u^{n+1} - v^{n+1} \right\rangle + \frac{1}{dt} \left\langle u^{n+1} - v^{n+1}, u^{n+1} - v^{n+1} \right\rangle -$$
$$\frac{1}{dt} \left\langle u^n - v^n, u^{n+1} - v^{n+1} \right\rangle = \left\langle -\hat{L}_h(u^n - v^n), u^{n+1} - v^{n+1} \right\rangle.$$

Rearranging terms and taking upper bounds we get the following:

$$\begin{aligned}
&\mu dt \left\langle p(u^{n+1}) - p(v^{n+1}), u^{n+1} - v^{n+1} \right\rangle + ||u^{n+1} - v^{n+1}||^2 \\
&\quad = \left\langle u^n - v^n, u^{n+1} - v^{n+1} \right\rangle + \left\langle -dt\hat{L}_h(u^n - v^n), u^{n+1} - v^{n+1} \right\rangle \\
&\quad = \left\langle (I - dt\hat{L}_h)(u^n - v^n), u^{n+1} - v^{n+1} \right\rangle \\
&\quad \leq ||(I - dt\hat{L}_h)(u^n - v^n)||||u^{n+1} - v^{n+1}|| \\
&\quad \leq ||(I - dt\hat{L}_h)||_{op}||u^n - v^n||||u^{n+1} - v^{n+1}||.
\end{aligned}$$

Note that $\mu dt \left\langle p(u^{n+1}) - p(v^{n+1}), u^{n+1} - v^{n+1} \right\rangle$ is non-negative by monotonicity of the subgradient of a convex function. We show this here by using the nonnegativity of Bregman distance:

$$\begin{aligned}
&0 \leq D_F^p(u^{n+1}, v^{n+1}) + D_F^p(v^{n+1}, u^{n+1}) \\
&\quad = F(u^{n+1}) - F(v^{n+1}) - \left\langle p(v^{n+1}), u^{n+1} - v^{n+1} \right\rangle + F(v^{n+1}) - F(u^{n+1}) - \left\langle p(u^{n+1}), v^{n+1} - u^{n+1} \right\rangle \\
&\quad = \left\langle p(u^{n+1}) - p(v^{n+1}), u^{n+1} - v^{n+1} \right\rangle.
\end{aligned}$$

Combining the positivity of the subgradient terms with equation above provides us with the following bound (assuming $||(I - dt\hat{L}_h)||_{op} \leq 1$):

$$||u^{n+1} - v^{n+1}|| \leq ||(I - dt\hat{L}_h)||_{op}||u^n - v^n|| \leq ||u^n - v^n||$$

as desired. □

Proof of theorem 2.5.2.

Proof. Considering the optimality condition for the energy (2.8) defining the implicit scheme, we see that the iterations for u^n and v^n can be written

$$\mu p(u^{n+1}) + \frac{u^{n+1} - u^n}{dt} = -\hat{L}_h u^{n+1} + f$$
$$\mu p(v^{n+1}) + \frac{v^{n+1} - v^n}{dt} = -\hat{L}_h v^{n+1} + f.$$

(If the operator $\hat{L}_h$ being considered in (2.8) is not positive semidefinite, then use (2.7) instead.) By taking the difference between these two equations we arrive at

$$\mu(p(u^{n+1}) - p(v^{n+1})) + \frac{1}{dt}(u^{n+1} - v^{n+1}) - \frac{1}{dt}(u^n - v^n) = -\hat{L}_h(u^{n+1} - v^{n+1}).$$

Next, taking the inner product of this equation with $u^{n+1} - v^{n+1}$ yields

$$\mu \left\langle p(u^{n+1}) - p(v^{n+1}), u^{n+1} - v^{n+1} \right\rangle + \frac{1}{dt} \left\langle u^{n+1} - v^{n+1}, u^{n+1} - v^{n+1} \right\rangle -$$
$$\frac{1}{dt} \left\langle u^n - v^n, u^{n+1} - v^{n+1} \right\rangle = \left\langle -\hat{L}_h(u^{n+1} - v^{n+1}), u^{n+1} - v^{n+1}. \right\rangle$$

Re-arranging terms and taking upper bounds we get the following:

$$\mu dt \left\langle p(u^{n+1}) - p(v^{n+1}), u^{n+1} - v^{n+1} \right\rangle + \left\| u^{n+1} - v^{n+1} \right\|^2$$
$$= \left\langle u^n - v^n, u^{n+1} - v^{n+1} \right\rangle + \left\langle -dt\hat{L}_h(u^{n+1} - v^{n+1}), u^{n+1} - v^{n+1} \right\rangle.$$

As in the explicit timestep case, $\langle p(u^{n+1}) - p(v^{n+1}), u^{n+1} - v^{n+1} \rangle \geq 0$ and so

$$||u^{n+1} - v^{n+1}||^2 \leq \left\langle u^n - v^n, u^{n+1} - v^{n+1} \right\rangle + \left\langle -dt\hat{L}_h(u^{n+1} - v^{n+1}), u^{n+1} - v^{n+1} \right\rangle.$$

If $\hat{L}_h$ is positive semidefinite then we have

$$||u^{n+1} - v^{n+1}||^2 \leq \left\langle u^n - v^n, u^{n+1} - v^{n+1} \right\rangle \leq ||u^n - v^n||||u^{n+1} - v^{n+1}||$$

and by canceling out terms we get the contractive inequality

$$||u^{n+1} - v^{n+1}|| \leq ||u^n - v^n||$$

as desired. □

Proof of Theorem 2.5.3:

Proof. We assume that S is stable in the following sense:

$$||\hat{u}^{n+1}|| \leq ||\hat{u}^n||$$

for some l^p norm; common choices would be the l^2 or l^∞ norms. Because the shrink operator decreases the magnitude of each component of a vector, it will (strictly, because $\mu > 0$) decrease whatever norm is chosen (in fact, the shrink operator is a contraction in all l^p norms). It follows easily that

$$||\hat{u}_\mu^{n+1}|| \leq ||Q(\hat{u}_\mu^n, \ldots, \hat{u}_\mu^{n-k})|| \leq ||\hat{u}_\mu^n||$$

so that the stability of S implies the stability of S_μ. In fact S_μ is more stable than S.

The key observation for showing consistency of S_μ is that while $\text{shrink}(\cdot, \mu)$ is nonlinear, the amount of this nonlinearity is bounded. In particular,

$$\text{shrink}(x, \mu) = x + O(\mu)$$

for any x, with $|O(\mu)| \leq \mu$. Applying this observation to the definition of the sparse scheme and assuming (for the purpose of local truncation analysis) that both schemes have the same starting points $\hat{u}_\mu^n = \hat{u}^n, \ldots, \hat{u}_\mu^{n-k} = \hat{u}^{n-k}$,

$$\begin{aligned}
\hat{u}_\mu^{n+1} &= \text{shrink}(Q(\hat{u}_\mu^n, \ldots, \hat{u}_\mu^{n-k}), \mu) \\
&= Q(\hat{u}_\mu^n, \ldots, \hat{u}_\mu^{n-k}) + O(\mu) \\
&= Q(\hat{u}^n, \ldots, \hat{u}^{n-k}) + O(\mu) \\
&= \hat{u}^{n+1} + O(\mu).
\end{aligned}$$

This shows that locally, S and S_μ differ only by a $O(\mu)$ quantity. This quantity may naively be accounted as part of the local truncation error for the sparse scheme, in which case

$$\tau_\mu^n = \tau^n + O(\mu)$$

where τ^n denotes the local truncation error of S and τ_μ^n the local truncation error of S_μ.

For the consistency of S_μ, we need the local truncation error to be greater than first order; assuming the consistency of S and that $\mu = O(dt^{1+\delta})$ yields this result. When $\mu = O(dt^p)$ for some p such that $\tau^n = O(dt^p)$ as well, $\tau_\mu^n = O(dt^p)$ and the order of convergence of the scheme is unchanged. □

Proof of Theorem 2.5.4:

Proof. First, recall that the optimality condition for (2.8) is

$$\mu p(\hat{u}_\mu^{n+1}) + (I + dt\hat{L}_h)\hat{u}_\mu^{n+1} - \hat{u}_\mu^n + dt\hat{f}_h = 0 \tag{2.16}$$

where $p(\hat{u}_\mu^{n+1}) \in \partial \left\|\hat{u}_\mu^{n+1}\right\|_1$. For simplicity of notation, let $w := (\hat{u}_\mu^{n+1} - \hat{u}^{n+1})$. Assuming (again for the purpose of local truncation analysis) that both schemes have the same starting point $\hat{u}_\mu^n = \hat{u}^n$, subtracting the ordinary backward Euler update from this gives

$$(I + dt\hat{L}_h)w = \mu p(\hat{u}_\mu^{n+1})$$

which implies

$$\left\|(I + dt\hat{L}_h)w\right\|_\infty \leq \mu.$$

Then, using the fact that $\hat{L}_h$ is positive definite, we get

$$\frac{\left\|(I + dt\hat{L}_h)w\right\|_\infty}{\|w\|_2 / N^{1/2}} \geq \frac{\left\|(I + dt\hat{L}_h)w\right\|_2}{\|w\|_2} \geq 1$$

which gives

$$\|w\|_{L^2(\Omega)} \sim \frac{\|w\|_2}{N^{1/2}} \leq \mu.$$

So, as with the explicit scheme,

$$\begin{aligned} & \hat{u}_\mu^{n+1} = \hat{u}^{n+1} + O(\mu) \qquad (\text{in } L^2(\Omega)) \\ \implies \quad & \tau_\mu^n = \tau^n + O(\mu) \end{aligned}$$

which yields consistency if $\mu = O(dt^{1+\delta})$ with $\delta > 0$, and implies the order of convergence is the same as that of the ordinary spectral scheme if $\mu = O(dt^p)$ with p such that $\tau^n = O(dt^p)$.

To prove stability of the scheme, return to (2.16) with $f = 0$ and take the inner product with $\hat{u}_\mu^{n+1}$ to get

$$\mu \left\|\hat{u}_\mu^{n+1}\right\|_1 + (\hat{u}_\mu^{n+1})^T (I + dt\hat{L}_h)\hat{u}_\mu^{n+1} - (\hat{u}_\mu^{n+1})^T \hat{u}_\mu^n = 0$$

which leads to

$$\begin{aligned}
\left\|\hat{u}_\mu^{n+1}\right\|_2^2 &\leq \langle \hat{u}_\mu^{n+1}, \hat{u}_\mu^n \rangle - \mu \left\|\hat{u}_\mu^{n+1}\right\|_1 - dt(\hat{u}_\mu^{n+1})^T \hat{L}_h \hat{u}_\mu^{n+1} \\
&\leq \langle \hat{u}_\mu^{n+1}, \hat{u}_\mu^n \rangle \\
&\leq \left\|\hat{u}_\mu^{n+1}\right\|_2 \left\|\hat{u}_\mu^n\right\|_2
\end{aligned}$$

and

$$\left\|\hat{u}_\mu^{n+1}\right\|_2 \leq \left\|\hat{u}_\mu^n\right\|_2,$$

as desired. □

Proof of theorem 2.6.2:

Proof. We have

$$\begin{aligned}
\frac{d}{dt}\frac{1}{2}||u(t,\cdot) - u_\delta(t,\cdot)||_2^2 &= \langle u - u_\delta, \partial_t u - \partial_t u_\delta \rangle \\
&= \langle u - u_\delta, -Lu + f - (-Lu_\delta + f - \delta\partial||\hat{u}(t)||_1) \rangle \\
&= -\langle u - u_\delta, Lu - Lu_\delta \rangle + \delta \langle u - u_\delta, \partial_0 ||\hat{u}_\delta||_1 \rangle \\
&\leq \delta \langle u - u_\delta, \partial_0 ||\hat{u}_\delta||_1 \rangle \\
&\leq \delta ||u - u_\epsilon||_2.
\end{aligned}$$

It follows that

$$\frac{d}{dt}||u(t,\cdot) - u_\epsilon(t,\cdot)||_2 \leq 2\delta$$

from which the result follows. □

Proof of theorem 2.8.1:

Proof. Let $\mathcal{F}_N[a(x/\epsilon)](k)$ denote the DFT of $a(x/\epsilon)$ on the grid; that is,

$$\mathcal{F}_N[a(x/\epsilon)](k) = \sum_{j=0}^{N-1} a\left(\frac{2\pi j}{N\epsilon}\right) e^{-2\pi i jk/N}.$$

Then

$$\begin{aligned}
\mathcal{F}_{2N}\left[a\left(\frac{x}{\epsilon/2}\right)\right](2k) &= \sum_{j=0}^{2N-1} a\left(\frac{2\pi j}{2N\cdot\epsilon/2}\right) e^{-2\pi i \frac{2k}{2N} j} \\
&= \sum_{j=0}^{N-1} a\left(\frac{2\pi j}{N\epsilon}\right)\left[e^{-2\pi i jk/N} + e^{-2\pi i(j+N)k/N}\right] \\
&= \sum_{j=0}^{N-1} a\left(\frac{2\pi j}{N\epsilon}\right)\left[e^{-2\pi i jk/N} + e^{-2\pi i jk/N} e^{-2\pi i k}\right] \\
&= 2\mathcal{F}_N[a(x/\epsilon)](k),
\end{aligned}$$

so that the even coefficients of $\mathcal{F}_{2N}\left[a\left(\frac{x}{\epsilon/2}\right)\right]$ are just those of $\mathcal{F}_N[a(x/\epsilon)]$. Also,

$$\begin{aligned}
\mathcal{F}_{2N}\left[a\left(\frac{x}{\epsilon/2}\right)\right](2k+1) &= \sum_{j=0}^{2N-1} a\left(\frac{2\pi j}{2N\cdot\epsilon/2}\right) e^{-2\pi i \frac{2k+1}{2N} j} \\
&= \sum_{j=0}^{N-1} a\left(\frac{2\pi j}{N\epsilon}\right)\left[e^{-2\pi i \frac{2k+1}{2N} j} + e^{-2\pi i \frac{2k+1}{2N}(j+N)}\right] \\
&= \sum_{j=0}^{N-1} a\left(\frac{2\pi j}{N\epsilon}\right) e^{-2\pi i \frac{2k+1}{2N} j}\left[1 + e^{-2\pi i \frac{2k+1}{2}}\right] \\
&= \sum_{j=0}^{N-1} a\left(\frac{2\pi j}{N\epsilon}\right) e^{-2\pi i \frac{2k+1}{2N} j}\left[1 + e^{-\pi i}\right] \\
&= 0
\end{aligned}$$

so that all odd coefficients vanish. These equalities give 2.15. □

References

[1] Eric Lewin Altschuler, Timothy J Williams, Edward R Ratner, Robert Tipton, Richard Stong, Farid Dowla, and Frederick Wooten. Possible global minimum lattice configurations for Thomson's problem of charges on a sphere. *Physical Review Letters*, 78:2681–2685, 1997.

[2] D Balagué, JA Carrillo, T Laurent, and G Raoul. Dimensionality of local minimizers of the interaction energy. *Archive for Rational Mechanics and Analysis*, 209(3):1055–1088, 2013.

[3] Andrew J Bernoff and Chad M Topaz. A primer of swarm equilibria. *SIAM Journal on Applied Dynamical Systems*, 10:212–250, 2011.

[4] Andrea L Bertozzi, José A Carrillo, and Thomas Laurent. Blow-up in multidimensional aggregation equations with mildly singular interaction kernels. *Nonlinearity*, 22(3):683, 2009.

[5] Andrea L Bertozzi and Thomas Laurent. Finite-time blow-up of solutions of an aggregation equation in r n. *Communications in mathematical physics*, 274(3):717–735, 2007.

[6] Andrea L Bertozzi, Thomas Laurent, and Flavien Léger. Aggregation and spreading via the newtonian potential: the dynamics of patch solutions. *Mathematical Models and Methods in Applied Sciences*, 22(supp01), 2012.

[7] Andrea L Bertozzi, Thomas Laurent, and Jesús Rosado. Lp theory for the multidimensional aggregation equation. *Communications on Pure and Applied Mathematics*, 64(1):45–83, 2011.

[8] H Brezis. OpeÌĄrateurs maximaux monotones et semigroupes de contractions dans les espaces de hilbert, 1973.

[9] Haim Brezis. *Monotone Operators Non Linear Semi-groups and Applications*. Université Pierre et Marie Curie, Laboratoire d'Analyse Numérique, 1974.

[10] Haim Brezis. Solutions with compact support of variational inequalities. *Russian Mathematical Surveys*, 29(2):103–108, 1974.

[11] Steven L Brunton, Jonathan H Tu, Ido Bright, and J Nathan Kutz. Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems. *arXiv preprint arXiv:1312.4221*, 2013.

[12] H. Cabral and D. Schmidt. Stability of relative equilibria in the problem of $n + 1$ vortices. *SIAM J. Math. Anal.*, 31:231–250, 1999.

[13] Russel E Caflisch. The fluid dynamic limit of the nonlinear boltzmann equation. *Communications on Pure and Applied Mathematics*, 33(5):651–666, 1980.

[14] Russel E Caflisch, Stanley J Osher, Hayden Schaeffer, and Giang Tran. Pdes with compressed solutions. *arXiv preprint arXiv:1311.5850*, 2013.

[15] Emmanuel J Candes, Yonina C Eldar, Thomas Strohmer, and Vladislav Voroninski. Phase retrieval via matrix completion. *SIAM Journal on Imaging Sciences*, 6(1):199–225, 2013.

[16] Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):11, 2011.

[17] Emmanuel J Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, 52(2):489–509, 2006.

[18] Emmanuel J Candès and Michael B Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.

[19] Rick Chartrand. Exact reconstruction of sparse signals via nonconvex minimization. *Signal Processing Letters, IEEE*, 14(10):707–710, 2007.

[20] Rick Chartrand. Shrinkage mappings and their induced penalty functions. In *IEEE Int. Conf. Acoust. Speech Signal Process*, 2014.

[21] Scott Shaobing Chen, David L Donoho, and Michael A Saunders. Atomic decomposition by basis pursuit. *SIAM journal on scientific computing*, 20(1):33–61, 1998.

[22] Yuxin Chen, Theodore Kolokolnikov, and Daniel Zhirov. Vortex swarms. *submittled*, 2013.

[23] Henry Cohn and Abhinav Kumar. Algorithmic design of self-assembling structures. *PNAS*, 106:9570–9575, 2009.

[24] C. Conca, E. Espejo, and K. Vilches. Remarks on the blowup and global existence for a two species chemotactic Keller–Segel system in R2. *European J. Appl. Math*, 22:553–580, 2011.

[25] Iain D. Couzin, Jens Krause, Nigel R. Franks, and Simon A. Levin. Effective leadership and decision-making in animal groups on the move. *Nature*, 433:513–516, 2005.

[26] Michael G Crandall and Thomas M Liggett. Generation of semi-groups of nonlinear transformations on general banach spaces. *American Journal of Mathematics*, pages 265–298, 1971.

[27] G. Crippa and M. Lécureux-Mercier. Existence and uniqueness of measure solutions for a system of continuity equations with non-local flow. *NoDEA: Nonlinear Differential Equations and Applications*, pages 1–15, 2011.

[28] Ingrid Daubechies, Olof Runborg, and Jing Zou. A sparse spectral method for homogenization multiscale problems. *Multiscale Modeling & Simulation*, 6(3):711–740, 2007.

[29] Marco Di Francesco and Simone Fagioli. Measure solutions for non-local interaction pdes with two species. *Nonlinearity*, 26(10):2777, 2013.

[30] David L Donoho. De-noising by soft-thresholding. *Information Theory, IEEE Transactions on*, 41(3):613–627, 1995.

[31] David L Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006.

[32] M.R. D'Orsogna, Y.L. Chuang, A.L. Bertozzi, and L.S. Chayes. Self-propelled particles with soft-core interactions: Patterns, stability, and collapse. *Phys. Rev. Lett.*, 96:104302, 2006.

[33] B. Düring, P. Markowich, J.F. Pietschmann, and M.T. Wolfram. Boltzmann and Fokker–Planck equations modelling opinion formation in the presence of strong leaders. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 465:3687–3708, 2009.

[34] Weinan E, Bjorn Engquist, and Zhongyi Huang. Heterogeneous multiscale method: a general methodology for multiscale modeling. *Physical Review B*, 67(9):092101, 2003.

[35] Yalchin Efendiev and Thomas Y Hou. *Multiscale finite element methods: theory and applications*, volume 4. Springer, 2009.

[36] C. Escudero, F. Macià, and J.J.L. Velázquez. Two-species-coagulation approach to consensus by group level interactions. *Physical Review E*, 82:016113, 2010.

[37] E.E. Espejo, A. Stevens, and J.J.L. Velázquez. Simultaneous finite time blow-up in a two-species model for chemotaxis. *Analysis*, 29:317–338, 2009.

[38] Lawrence C Evans. *Partial differential equations*. Providence, Rhode Land: American Mathematical Society, 1998.

[39] R.C. Fetecau, Y. Huang, and T. Kolokolnikov. Swarm dynamics and equilibria for a nonlocal aggregation model. *Nonlinearity*, 24:2681, 2011.

[40] Tom Goldstein and Stanley Osher. The split bregman method for l1-regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009.

[41] Gene H Golub and Charles F Van Loan. *Matrix computations*, volume 3. JHU Press, 2012.

[42] J. M. Haile. *Molecular Dynamics Simulation: Elementary Methods*. John Wiley & Sons, Inc., New York, NY, USA, 1st edition, 1992.

[43] Trevor Hastie, Robert Tibshirani, Jerome Friedman, T Hastie, J Friedman, and R Tibshirani. *The elements of statistical learning*. Springer, 2009.

[44] Darryl D. Holm and Vakhtang Putkaradze. Aggregation of finite-size particles with variable mobility. *Phys. Rev. Lett.*, 95:226106, 2005.

[45] D. Horstmann. Generalizing the Keller–Segel model: Lyapunov functionals, steady state analysis, and blow-up results for multi-species chemotaxis models in the presence of attraction and repulsion between competitive interacting species. *Journal of Nonlinear Science*, 21:231–270, 2011.

[46] Thomas Y Hou, Zuoqiang Shi, and Peyman Tavallali. Sparse time frequency representations and dynamical systems. *arXiv preprint arXiv:1312.0202*, 2013.

[47] Thomas Y Hou and Xiao-Hui Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of computational physics*, 134(1):169–189, 1997.

[48] J. H. Irving and John G. Kirkwood. The statistical mechanical theory of transport processes. iv. the equations of hydrodynamics. *The Journal of Chemical Physics*, 18:817–829, June 1950.

[49] Ioannis G Kevrekidis, C William Gear, James M Hyman, Panagiotis G Kevrekidid, Olof Runborg, Constantinos Theodoropoulos, et al. Equation-free, coarse-grained multiscale computation: Enabling mocroscopic simulators to perform system-level analysis. *Communications in Mathematical Sciences*, 1(4):715–762, 2003.

[50] T. Kolokolnikov, Y. Huang, and M. Pavlovski. Singular patterns for an aggregation model with a confining potential. *Physica D*, 2012.

[51] Theodore Kolokolnikov, Hui Sun, David Uminsky, and Andrea L. Bertozzi. Stability of ring patterns arising from two-dimensional particle interactions. *Phys. Rev. E*, 84:015203, 2011.

[52] Tijana Kostić and Andrea Bertozzi. Statistical density estimation using threshold dynamics for geometric motion. *J Sci Comput*, 54:513–530, 2013.

[53] Thomas Laurent. Local and global existence for an aggregation equation. *Communications in Partial Differential Equations*, 32(12):1941–1964, 2007.

[54] Herbert Levine, Eshel Ben-Jacob, Inon Cohen, and Wouter-Jan Rappel. Swarming patterns in microorganisms: Some new modeling results. *Proceedings IEEE CDC*, 2006.

[55] Herbert Levine, Wouter-Jan Rappel, and Inon Cohen. Self-organization in systems of self-propelled particles. *Physical Review E*, 63:017101, 2000.

[56] Yao li Chuang, Maria R. D'Orsogna, Daniel Marthaler, Andrea L. Bertozzi, and Lincoln S. Chayes. State transitions and the continuum limit for a 2D interacting, self-propelled particle system. *Physica D*, 232:33–47, 2007.

[57] Yao li Chuang, Yuan R. Huang, Maria R. D'Orsogna, and Andrea L. Bertozzi. Multi-vehicle flocking: Scalability of cooperative control algorithms using pairwise potentials. *2007 IEEE International Conference on Robotics and Automation*, 2007.

[58] Tianbo Liu. Hydrophilic macroionic solutions: What happens when soluble ions reach the size of nanometer scale? *Langmuir*, 26:9202–9213, 2009.

[59] Tianbo Liu, Melissa LK Langston, Dong Li, Joseph M Pigga, Céline Pichon, Ana Maria Todea, and Achim Müller. Self-recognition among different polyprotic macroions during assembly processes in dilute solution. *Science*, 331:1590–1592, 2011.

[60] Alan Mackey, Theodore Kolokolnikov, and Andrea Bertozzi. Two-species particle aggregation and stability of co-dimension one solutions. *Discrete and Continuous Dynamical Systems*, 34:1411–1436, 2014.

[61] Alan Mackey, Hayden Schaeffer, and Stanley Osher. On the compressive spectral method. *Multiscale Modeling and Simulation*, 12(4):1800–1827, 2014.

[62] Andrew J. Majda and Andrea L. Bertozzi. *Vorticity and Incompressible Flow.* Cambridge University Press, 2002.

[63] Alex Mogilner, L Edelstein-Keshet, L Bent, and A Spiros. Mutual interactions, potentials, and individual distance in a social aggregation. *Journal of mathematical biology*, 47:353–389, 2003.

[64] Balas Kausik Natarajan. Sparse approximate solutions to linear systems. *SIAM journal on computing*, 24(2):227–234, 1995.

[65] Lance J Nelson, Gus LW Hart, Fei Zhou, and Vidvuds Ozoliņš. Compressive sensing as a paradigm for building physics models. *Physical Review B*, 87(3):035125, 2013.

[66] James Nolen, George Papanicolaou, and Olivier Pironneau. A framework for adaptive multiscale methods for elliptic problems. *Multiscale Modeling & Simulation*, 7(1):171–196, 2008.

[67] Bruno A Olshausen and David J Field. Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision research*, 37(23):3311–3325, 1997.

[68] Vidvuds Ozoliņš, Rongjie Lai, Russel Caflisch, and Stanley Osher. Compressed modes for variational problems in mathematics and physics. *Proceedings of the National Academy of Sciences*, 110(46):18368–18373, 2013.

[69] Vidvuds Ozoliņš, Rongjie Lai, Russel Caflisch, and Stanley Osher. Compressed plane waves yield a compactly supported multiresolution basis for the laplace operator. *Proceedings of the National Academy of Sciences*, 111(5):1691–1696, 2014.

[70] G Papanicolau, A Bensoussan, and J-L Lions. *Asymptotic analysis for periodic structures*. North Holland, 1978.

[71] Rosa Ramírez and Thorsten Pöschel. Coefficient of restitution of colliding viscoelastic spheres. *Phys. Rev. E*, 60:4465, 1999.

[72] Benjamin Recht, Maryam Fazel, and Pablo A Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, 52(3):471–501, 2010.

[73] M.C. Rinaldo, M. Garavello, and M. Lécureux-Mercier. A class of nonlocal models for pedestrian traffic. *Mathematical Models and Methods in Applied Sciences*, 22:1150023, 2012.

[74] P. Romanczuk, M. Bar, W. Ebeling, B. Lindner, and L. Schimansky-Geier. Active brownian particles: From individual to collective stochastic dynamics. *Eur. Phys. J. Special Topics*, 202:1–162, 2012.

[75] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[76] Hayden Schaeffer, Russel Caflisch, Cory D Hauck, and Stanley Osher. Sparse dynamics for partial differential equations. *Proceedings of the National Academy of Sciences*, 110(17):6634–6639, 2013.

[77] Jean-Luc Starck, Michael Elad, and David Donoho. Redundant multiscale transforms and their application for morphological component separation. *Advances in Imaging and Electron Physics*, 132(82):287–348, 2004.

[78] Hui Sun, David Uminsky, and Andrea L. Bertozzi. A generalized Birkhoff-Rott equation for two-dimensional active scalar problems. *SIAM J. Appl. Math*, 72:382–404, 2012.

[79] Hui Sun, James von Brecht, David Uminsky, Theodore Kolokolnikov, and Andrea Bertozzi. Ring patterns and their bifurcations in a nonlocal model of biological swarms. *preprint*, 2012.

[80] JI Tello and M. Winkler. Stabilization in a two-species chemotaxis system with a logistic source. *Nonlinearity*, 25:1413, 2012.

[81] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[82] Chad M Topaz, Andrew J Bernoff, Sheldon Logan, and Wyatt Toolson. A model for rolling swarms of locusts. *The European Physical Journal Special Topics*, 157:93–109, 2008.

[83] Chad M Topaz, Andrea L Bertozzi, and Mark E Lewis. A nonlocal continuum model for biological aggregations. *Bulletin of Mathematical Biology*, 68:1601–1623, 2006.

[84] Joel A Tropp and Anna C Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *Information Theory, IEEE Transactions on*, 53(12):4655–4666, 2007.

[85] Lev Tsimring, Herbert Levine, Igor Aranson, Eshel Ben-Jacob, Inon Cohen, Ofer Shochet, and William N Reynolds. Aggregation patterns in stressed bacteria. *Physical review letters*, 75:1859–1862, 1995.

[86] James H von Brecht and David Uminsky. On soccer balls and linearized inverse statistical mechanics. *Journal of Nonlinear Science*, 22:935–959, 2012.

[87] James H von Brecht, David Uminsky, Theodore Kolokolnikov, and Andrea L Bertozzi. Predicting pattern formation in particle interactions. *Mathematical Models and Methods in Applied Sciences*, 22, 2012.

[88] Wikipedia. Finite element method — wikipedia, the free encyclopedia. `http://en.wikipedia.org/w/index.php?title=Finite_element_method&oldid=639712628`, 2014. [Online; accessed 16-January-2015].

[89] Wikipedia. Heap (data structure) — wikipedia, the free encyclopedia. `http://en.wikipedia.org/w/index.php?title=Heap_(data_structure)&oldid=641306365`, 2015. [Online; accessed 16-January-2015].

[90] G. Wolansky. Multi-components chemotactic system in the absence of conflicts. *European J. Appl. Math*, 13:641–661, 2002.

[91] John Wright, Allen Y Yang, Arvind Ganesh, Shankar S Sastry, and Yi Ma. Robust face recognition via sparse representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(2):210–227, 2009.